AGARD CONFERENCE PROCEEDINGS No. 57

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92 NEUILLY SUR SEINE FRANCE

Problems in Mechanization of Small Information Centres

NORTH ATLANTIC TREATY ORGANIZATION

INITIAL DISTRIBUTION IS LIMITED
FOR ADDITIONAL COPIES SEE BACK COVER

131

AGARD Conference Proceedings No. 57

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

PROBLEMS IN MECHANIZATION OF SMALL INFORMATION CENTRES

Papers presented at the Specialist Meeting of the Technical Information Panel of AGARD,
in Ottawa, Canada, 16-17 September 1969

ACKNOWLEDGMENTS

Mr J.J.Irvine is thanked for his help in compiling the discussions which followed the presentation of the papers in due time for their inclusion in these proceedings.

The authors of papers and other contributors are also thanked for the prompt provision of their material in suitable form for reproduction.

Published February 1970

659.2:681.3.01:025.5

Printed by Technical Editing and Reproduction Ltd
Harford House, 7-9 Charlotte St, London W1P 1HD

SUMMARY

This volume contains the text of eleven papers presented at the Specialist Meeting of the Technical Information Panel of AGARD, which was held at the National Library and Archives Building in Ottawa on 16th and 17th September 1969.

The purpose of this meeting was to discuss and exchange experiences concerning the many practical problems that occur in the setting up of an automated centre; from the seeking of approval for the project, to the commencement of its successful operation.

Topics covered are: the advantages to be gained with the establishment and operation of Small Information Centres, the role of Small Information Centres in the NATO community, assessing the value of the pilot project approach, the importance of greater participation by the user both in the preparation of better profiles and in the provision of feedback data, the comparison of cost and effectiveness of the operation, assessment and selection of hardware, the importance of training and retraining of personnel.

PREFACE

In recent years the Technical Information Panel of AGARD has devoted a great deal of attention to the automation of information systems. In particular, the experience gained in this field by the large United States agencies and European organizations has been studied and discussed in great detail. Smaller centers are now being developed and operated in a number of the member countries. The purpose of this Specialist Meeting was to discuss and exchange experiences in the many practical problems that occur from the seeking of approval for an automated center to the commencement of its successful operation as a production unit. It was intended to cover the whole spectrum of problem types encountered in setting up an automated center, in a sequence in which they might logically be expected to occur.

Some of the important aspects covered are:

(a) The advantages to be gained with the establishment and operation of small information centers.

(b) The role of Small Information Centers in the NATO community.

(c) Assessing the value of the pilot project approach.

(d) The importance of greater participation by the user both in the preparation of better profiles and feedback data.

(e) Comparing cost and effectiveness of the operation.

(f) Assessment and selection of hardware.

(g) The importance of training and retraining of personnel in order to remain current with the latest "State of the Art".

It was quite evident from the comments received, both verbal and by correspondence, that a coherent picture was presented and benefited the majority of the attendees at the meeting.

CONTENTS

OPENING REMARKS

R.L.Marderosian

Deputy Chairman, Technical Information Panel AGARD,
Wright-Patterson AFB, Ohio, USA

Distinguished guests, ladies and gentlemen -

It gives me great pleasure to welcome all of you on behalf of the Technical Information Panel to this Specialist Meeting on "Problems in Mechanization of Small Information Centres" that will continue for the next two days. You will hear presentations from some of the most noted personnel in the field of information science from the various countries of NATO. We encourage audience participation in the form of questions, enquiries, recommendations, and discussions after the completion of each paper.

(Mr Marderosian then introduced the first speaker in the Opening Session, Mr G.Sylvestre, who was followed by Mr F.R. Thurston and Mr H.A.Stolk.)

A WELCOME TO THE NATIONAL LIBRARY

Guy Sylvestre

National Librarian, Ottawa, Canada

Mr Chairman,
Ladies and gentlemen,

I welcome you all most warmly to the National Library of Canada. I was delighted to hear sometime ago that AGARD had chosen to hold this meeting of specialists in this institution, and I sincerely hope that you too will be delighted to have come here and that, when you go back home, you will keep many happy memories of your visit to Ottawa and of your meetings here which, I trust, will be thoroughly successful, thoroughly fruitful, gratifying and enjoyable.

When I looked at your programme, I found that you will be addressing yourselves to problems which very much exercise our minds these days. The National Library is not exactly a small information centre, but we do have yet to solve some of the problems which you will discuss here today and tomorrow. We now have under way a comprehensive study, encompassing the potential benefits of electronic data processing in all areas of activity, with a view to establishing an integrated information system in the Library. This is not meant only to assist us in improving our own housekeeping operations, but it is hoped that, in cooperation with other large research libraries in Canada, with which we exchange technical information, there will be established in this country an integrated network of libraries providing for the fullest interchange of information through the achievement of a large measure of compatibility of processing schemes, rules and formats. Such exchanges of information are also taking place with sister-institutions in several countries, for we are all very much concerned about the problems of inputting data from other systems, as we librarians have long been concerned with achieving greater uniformity of cataloguing rules and of indexing periodical literature with our traditional manual methods. As you all very well know, compatibility, when it can be achieved, is even more desirable among automated systems than between manual ones and it is gratifying to note that, although the situation is far from ideal and if much remains to be done to achieve a larger degree of uniformity among information systems, there is a growing recognition, a growing awareness in the library world today of the magnitude and the complexity of the problems involved and, it seems to me, a genuine desire, I would even say a deep eagerness, on the part of many, to cooperate more closely than ever before in order to break barriers and to facilitate and expedite the interchange of information on a world-wide basis. I am very optimistic, and your presence here, ladies and gentlemen, gives evidence of this readiness and makes me even more optimistic in this respect.

Je ne voudrais pas, mesdames et messieurs, que vous croyiez que, si je vous ai indiqué brièvement comment l'avenir nous préoccupait, c'est que nous n'ayions rien fait jusqu'ici. Nous croyons, en effet, avoir réussi pas mal en très peu de temps, ce qui ne veut pas dire que nos projets ne sont pas plus ambitieux. Il ne faut pas oublier que cette Bibliothèque nationale n'existe que depuis 1953 et que nous ne sommes dans nos meubles que depuis deux ans. Nous publions déjà, cependant, une bibliographie nationale qui est, je crois, une des meilleures qui soient, et nous avons publié d'utiles répertoires de revues et de journaux, ainsi que des bibliographies. Nous avons aussi établi un catalogue collectif national qui comprend plus de 10.000.000 de fiches représentant les collections de plus de 300 bibliothèques canadiennes; qui grandit au rythme actuel de 4.700 fiches par jour et qui nous permet de localiser rapidement les ouvrages dont les chercheurs ont besoin. Ainsi, l'an dernier, on nous a demandé de localiser 81.373 titres - soit une moyenne de 373 par jour - et nous en avons trouvé plus de 64.000 au Canada, soit 80 pour cent. La moitié de ces demandes nous parviennent par Telex - 60 bibliothèques canadiennes communiquent avec nous par ce moyen - et nous leur répondons de la même manière, ce qui facilite et accélère singulièrement les travaux de nombreux chercheurs dans toutes les parties de ce grand pays qu'est le nôtre. (Incidemment, nous sommes aussi reliés aux grandes bibliothèques américaines et étrangères par TWX.)

En contribuant ainsi à mettre rapidement à la disposition des chercheurs les ressources des principales bibliothèques du pays, la Bibliothèque nationale du Canada est pour ainsi dire le coeur d'un réseau dont les battements font circuler le sang jusqu'aux membres les plus extrêmes, et qui permet aussi à ces membres de se vivifier réciproquement. Cette osmose féconde fait pour ainsi dire des bibliothèques des vases communicants, ce qui contribue à élever le niveau de leurs services respectifs.

Je m'en excuse, monsieur le président, j'ai déjà été trop long, mais j'ai cru que vous seriez peut-être intéressé à avoir quelque idée de ce que nous faisons. Je vous souhaite la bienvenue, je formule le voeu que vos travaux soient aussi fructueux que votre séjour parmi nous, agréable.

ADDRESS OF WELCOME

F.R. Thurston

AGARD National Delegate
Director, National Aeronautical Establishment, Ottawa, Canada

Mr Chairman, Ladies and Gentlemen, Panel Members, Distinguished Visitors –

I am delighted to have this opportunity to welcome the Members of the Technical Information Panel to Canada. I understand that there are representatives from France, Germany, The Netherlands, Portugal, Norway, U.K., U.S.A. and Italy.

The last time Canada was privileged to be host to the technical information group of AGARD was in June 1955. At that time it was known as the Technical Information and Documentation Committee. However, at the beginning of 1965 the Committee status was raised to that of Panel. This raised status gives some indication of the increasing importance of the role of the TIP in AGARD and, to a substantial extent, in NATO itself.

We are particularly happy to host this meeting and I am sure the Canadian members of the Panel are as pleased as I am, to have this opportunity to return, in some measure, the generous hospitality we have enjoyed in many other NATO countries.

May I also say how pleased I am to be able to welcome so many colleagues from Europe. Canada is favoured by being a close neighbour of the United States which makes the exchange of visits and ideas simple and frequent. However, Canada is too far from Europe to afford many such interchanges. Consequently international meetings of this nature are particularly welcome in Canada and it is my hope that our European friends will benefit as much from being in Canada as we will from having them with us.

It is a privilege too, on this occasion, to be able to hear a number of speakers from abroad present their various approaches to problems in information science and to learn of the successes achieved and to be warned against approaches which have proved fruitless.

The achievements of the Technical Information Panel have been many and its function is such that it has influenced the whole of AGARD. For example, many of us who are familiar with AGARD reports and other AGARD publications may not have been aware that the Technical Information Panel has played a major role in the publications programme, and contrived the distribution of these publications through national centres in member countries.

It is gratifying to see, from the topics to be presented today and tomorrow, that the Panel has not been content to restrict itself to traditional methods of handling information which are not now adequate to deal with the cataract of technical material flowing out of the world's research laboratories.

Much more could be said about the Panel. I will, however, leave it to the Chairman to tell you of its aims, accomplishments and plans.

I offer you my best wishes for the success of the Specialist Meeting here in the National Library. I also hope that the Members of the Panel will enjoy the facilities of the National Research Council when they hold their business meeting there later in the week.

KEYNOTE ADDRESS

H. A. Stolk

Chairman, Technical Information Panel AGARD,
National Aerospace Laboratory (NLR), Amsterdam, The Netherlands

After a short presentation on the organization and mission of the Advisory Group for Aerospace Research and Development, a NATO Agency which was instituted in 1951 on the recommendation of the late Prof. Theodore von Kármán, and in which the leading personalities of the NATO nations in the fields of aerospace science and technology are brought together, Mr Stolk proceeded:

The present Technical Information Panel was formed in 1953 as a Documentation Committee, to provide service and advice to NATO in the documentation field and to advise the AGARD Panels on publication problems. With the great increase in technical and scientific information, the Committee had considerable success in encouraging national documentation agencies to achieve better exchange of information, and to make it available to larger numbers of NATO research workers. The problem of improving the exchange of information in the defence field was and still is a major topic of discussion. Though the recommendations of the Panel to this effect laid down in a report in 1962 were subsequently approved by the NATO Council, their implementation could only partly be achieved, security regulations and proprietary rights being the greatest obstacle and will be so, for a long time to come. The panel being the only permanent organisation within NATO concerned with procedures and operational methods in the information field, still some progress could be made. The fact that the Directors and senior officers of most of the National Defense Documentation Centres are personnel members greatly facilitated the communication on these problems. Panel meetings formed a favourable occasion to get to know each other, to exchange experience which often led to considerable improvements of methods and organisation. The different problems being recognised permitted to define them better and in some cases to attenuate or solve them. They also have contributed to a direct exchange of documentary information. The different heads of the armed forces information services speak about the same language and have obtained a sort of common view of their work. The conclusions from panel discussions have permitted their centres to better organise the collection of national documents and consequently to make them available to the NATO community as a whole. For your information I should like to dwell a moment on the duties of an armed forces information centre as its activities are so very closely related to those of TIP. Probably the most important duty is the processing of unpublished reports. This includes the circulation of reports from the country's own R&D establishments and the collection and circulation of reports from abroad. Instructions on circulation have to be obtained from the appropriate authority and, unless the circulation is unlimited, a close check has to be kept on the security and other restrictions which control dissemination. All the reports must be recorded and indexed and with a few exceptions are announced generally in an abstract journal, which has a wide but controlled distribution to those working in the field. Requests for copies of reports which are subject to certain limitations must be dealt with on a need-to-know basis.

As has been said before, the strength of the Panel is largely due to the support given by the heads of various centres who have done much to improve the exchanges of report literature between NATO countries. The value of personal contact between the staff of the various national centres can hardly be underestimated and this has done much to increase efficiency. Because the centres deal with the whole of defence literature, a very vast field indeed, the Panel is able to assist NATO both in and outside the AGARD limitations of aerospace. The centres thus can be an extremely important source of information. Though their main efforts are available to those related to defence, service will in many cases be provided to those outside this field.

To speak of the work of the Panel, it is necessary to be selective and the following items will probably be of interest: The Panel has assisted three countries, that is Turkey, Greece and Portugal, which were interested in setting up documentation centres by appointing TIP specialists to visit, determine the local problems and advise on the scale and organisation of the unit required. Especially the efforts by Turkey and Portugal have made a great deal of progress.

An examination of the possibility of standardising documentary procedure has had limited success and is continuing. Little difficulty is experienced at present by lack of uniformity but the spread of mechanization offers opportunities which can only be fully realized by standardisation. This gives an opportunity to solve the problems of methods and their compatibility. Think only of the exchange of magnetic tapes, of the utilisation of the same vocabulary, and you will see how this could facilitate the transfer of information.

Bibliographies, listing the published and unpublished literature in certain technical fields, have been prepared for AGARD technical panels. Most of the national members of TIP have contributed, so the coverage is better than could be achieved by any single person or organisation.

The Panel prepared and produced the AGARD multi-lingual dictionary of aeronautics in nine languages and a glossary of terms on documentation.

Over the years the Panel has discussed storage and retrieval of information and in 1968 a joint symposium was held in Munich with the Avionics Panel to provoke a user-supplier dialogue in this field. In fact, the symposium was held principally for the benefit of scientists and engineers who use information services and who would participate in a user-supplier dialogue on the present state of the art and the basic design of future systems The main object was not fully achieved however. Either the users were not present in any considerable numbers or if present they did not contribute very much to the discussion. To pursue the dialogue and to inform a still greater number of users in NATO of the present possibilities and those in the near future, a lecture series to be held in 1970 was proposed and approved by the NDB. Invitations thus far have been received from Norway and Italy. The lecture series is intended to give the users an overall briefing of the situation in the field of information and to provide at the same time a forum for a broad exchange of views between information people, scientists and engineers.

Though the difficulties to get something accomplished internationally are great, the panel nevertheless feels that its efforts have already contributed much to the solution of many common problems in the various NATO countries in the defence documentation field.

PAPER 1

OBTAINING MANAGEMENT ACCEPTANCE OF A SYSTEM PROPOSAL

by

A.C. Jones

Director, Defence Scientific Information Services,
Defence Research Board, Ottawa, Canada

SUMMARY

The paper suggests the steps necessary to prepare an automated information retrieval system proposal, to obtain data for justification of the proposal and to present it for final approval. The need is stated for the information centre to be itself convinced that it has the best proposal to offer, before it can legitimately expect to receive approval. The higher management may need to be shown what the information centre has been doing and what its customers want, before examining the new proposal. The paper touches on the difficult question of evaluating any information system, but emphasises the impracticability of retaining a manual system when the raw material for processing is going to be provided in machineable form. Finally, it urges a team approach to preparing and presenting the proposal, and holding the team together for follow-up on points requiring further resolution.

SOMMAIRE

L'auteur indique, à titre de suggestion, les démarches qui s'imposent lorsqu'on veut préparer une proposition de système de recherche d'information automatisé, obtenir les données justifiant cette proposition, et la présenter pour approbation définitive. Il souligne la nécessité, pour le centre d'informations, d'être lui-même convaincu que la proposition qu'il soumet est la meilleure, avant de s'attendre légitimement à ce qu'elle soit acceptée. Avant d'examiner la nouvelle proposition, il se peut que les cadres supérieurs de la direction aient besoin de savoir ce que le centre d'information a réalisé, et ce que désirent ses utilisateurs. L'auteur aborde le problème délicat que soulève l'évaluation de tout système d'informations, mais souligne l'impossibilité de garder un système manuel lorsque la matière première à traiter doit être fournie sous une forme propre à un traitement mécanique. Enfin, il recommande vivement la constitution d'une équipe pour préparer et présenter la proposition, et le maintien de cette équipe pour l'étude des problèmes ultérieurs.

OBTAINING MANAGEMENT ACCEPTANCE OF A SYSTEM PROPOSAL

A.C. Jones

INTRODUCTION

Obtaining the formal approval of your higher management is not of course the first thing you have to do to get a new information system into being, but it is the most vital step. It has to be preceded by a good deal of work to prepare the system proposal. This work either requires extra money and help, or else diverts these from the normal activity of your information centre. In either case you will need some form of approval in principle, or at least sympathetic understanding from your management and your customers while you are doing it. The work itself must be carried out with two objectives, one to devise the best system that seems feasible, the other to present the proposal in such a way that it is accepted by your organisation. In other words, you must make a good plan, and you must sell it. My assignment is to talk about the selling process, and from now on the objective I shall be referring to is that of selling the proposal.

We in DSIS have been going through the system development process during the past three years, and the experience we have had is the basis for this talk. Circumstances alter cases, so the best I can hope for is that some of the lessons we learned may be useful to you. They are offered in that spirit.

GENERAL EXPERIENCE

You are not going to hear any great secrets. We made no great discoveries, but just learned once again that there is no substitute for hard, painstaking preparatory work, and that this work must be planned to meet its objective.

Are there any general guides to follow? There is certainly one we could have learned from any good salesman, namely, that you must have confidence in your product before you can successfully and honestly sell it to others. You must be completely satisfied with your proposal before you put it up for approval or you may either not sell it, or perhaps live to regret that you did. Your management cannot re-examine all the detailed studies you will have made in formulating your proposal. They will have to take most of it on faith - faith in you! ·u should be your own severest critic.

A second general guide we could have learned from any military staff officer. Get all the intelligence you can about your target. First, what does your management need to have in any proposal; what does it like to have; how and when is a proposal best presented? What kind of questions are usually asked after a presentation, by which members of the management with what interests in mind? When you have collected answers to these, and other similar questions that will quickly come to mind in your own organisation, examine every facet of your proposal in their light. This should tell you which points will score and which ones would miss the target.

APPROACH TO MANAGEMENT

As we now see it, the rest is a function of detailed analysis of the problem itself, of possible solutions and of presenting the solution you have selected. But first, does your management really know what you have to do now? Do they know who your users are, what special needs they have, how many different kinds and forms of information have to be handled? Do they know what methods you already use, what skills your personnel have to have, what complexities are inherent in the present system?

Probably not. In any organisation with a technical function such as ours, the bulk of the information we .andle is technical, and higher management is one place where no one has much time available to keep up with technical literature. So management may be the centre's poorest customer, and therefore won't know the present system from direct experience.

Can they be expected to fully appreciate your new proposal under these circumstances? It would be unrealistic to expect this, and unfair to your proposal to hazard it in such a gamble. You must try to educate them. Perhaps you can do this at the same time as you present your proposal. We decided against doing that - it seemed too big a bite with too little time for digesting it. We were fortunate in being granted two opportunities to brief our management, several weeks apart, and several months before presenting our full proposal (a four-year plan). We told them what volume and k'·ds of material we handled, what volume of material was involved in each of the regular services we were providing. We tried to devise measures of the effort that had to go into meeting various requests for searches, and then presented data based on these measurements. We showed the relative extent to which the various services were used by customers in different categories, the

categories being chosen in terms meaningful to management, that is administratively meaningful. We showed how our personnel were deployed, and how we used the money presently allotted. We were lucky to be able to borrow an operational researcher with experience in studying and costing flow processes. He did a complete study of our operation, which provided a basis for that aspect of our briefing to management.

All this material will have taken many weeks or months to compile. It will almost certainly pay off. Management would have asked most of those questions anyhow, and your credibility will be much higher if you already have the answers. Incidentally, we learned a great deal more about our own operations by this means, and discovered some ways of improving them.

The products of information work, an announcement, an SDI service, a bibliography, a few reports supplied against a search request, seem quite simple in themselves. Higher management doesn't necessarily understand how complex are the processes that produce these end results, nor the intellectual effort that goes into them. It is worth a major effort in the briefing process to get this idea across. We tried various ways such as flow charts, breakdowns of job descriptions, and of work assignments. Individually, none of these seemed very effective, collectively they may have had some impact.

The organisations in which our information centres operate have many reasons for their existence. Service to a body of customers is often quite a small or even a non-existent element of their mission. The higher management therefore, may not be service-oriented. The information centre has to be - there is no justification for its existence except service to its users. Implicit in a service operation is the guarantee of maintenance of high standards of quality, timeliness, and overall reliability. Other kinds of operation can vary their output to suit their internal needs. A service operation must provide its output to suit its customers' needs. If you can get this point clearly established with your management, you will save many future pains. We tried to stress this in briefings, but we have found that our customers can make this point more effectively than we can, provided they direct their complaints to the management and not to the information centre.

THE MAIN PROBLEM

Management now knows what you are doing, and why. The next step is to explain why you must change a lot of this. Everyone has heard of the information explosion, but facts and figures to show how it affects your own information centre are necessary. We would have liked to present at least two aspects; what does the steady growth in volume of relatively unselected information cost our customers in terms of the effort they have to expend on information searching; what would continuation of our present system mean in terms of annual increase of information staff and money to handle the increasing volume? These two considerations should give some indication of the scope of change which the organisation could afford to consider, in respect of allocation of resources to a new information system. We were unable to do very much at the time to assess the burden on the customers, but greater efforts to study this might be profitable. We were of course able to provide a forecast of our own internal burden and its growth rate.

One further and very important consideration - can we really go on using old methods? Not only is the volume of information increasing but the form in which it is available is changing. If you can handle it directly in these newer forms, there is a substantial saving in effort to be had. Soon you may be unable to get much of the information in paper form, and it will cost extra time and money to convert it back into paper so that it can be handled by the old methods.

Thus there are two major selling points to emphasise:

(a) the cost is going up even if you stick to the old system;

(b) the old system is going to become impracticable to maintain.

THE CHOICE OF SOLUTION

You will have many things to consider, but hopefully you will have a good deal of expert help in deciding what to propose as a solution to the problem. My only contribution is to suggest how it might be presented. The previous indoctrination should have paved the way for the proposal. If necessary, indoctrination should have been held back until the form the proposal would take was sufficiently clear that it could fulfil that purpose.

The proposal will describe the problem and the nature of the solution proposed. Beyond the obvious requirement for this to be logical, reasonable, and understandable, the main question for the management will be "can we afford it?".

We put the proposal in the form of resources wanted, manpower and money, and when required. It clearly indicated the total resources required, as well as annual amounts, and showed what existing resources would be phased out as they were replaced by elements of the new system. The terms used to explain this should allow management to weigh up your demands against those from other branches of the organisation, without frustrating difficulties. In our case, management had recently instituted a procedure to do this, and we merely followed that procedure.

The presentation must show what the organisation is to get out of it. We shouldn't try to claim that great savings will accrue - they won't. It is difficult to present valid cost-benefit arguments, because it is still impossible to place a money value on information work as a whole. An occasional case, yes. Generally, no. It is possible to demonstrate cost-effectiveness. The new system will be able to handle more jobs, do them faster with less drudgery, do many jobs the old system couldn't handle at all, do everything more frequently, and provide data for process control without the costly maintenance of statistics by clerical means.

In other words it is possible to show that -

(a) you will be able to beat the explosion problem,

(b) you will be able to give a better service without significantly greater demands on future resources than you would have had to make if the old system were kept.

MECHANICS OF PRESENTATION

Just a few words on handling the presentation. You will have accumulated lots of data, but pick out only the significant material. Present this visually if possible, but only if good quality reproduction is available. Use a team approach if you can. Let the people who helped prepare the proposal share in its presentation. You will do a better job of presenting your case, and you will have the experts with you to answer the searching questions - better answers than you might have given on your own. Keep the presentation as short as you can. Your management have lots of other problems to consider, and it won't help your case to try to oversell it. Have your complete presentation available in written and tabular form as well. It is almost sure to be needed in both forms at some time.

FINAL CONSIDERATION

A final word of warning. Don't sit back and wait for the big day of the presentation. It is unlikely that you will get either a yes or no answer. A more likely outcome is a conditional yes; that is, the plan seems acceptable, but would you please examine certain points and provide a supplementary brief. If you could guess in advance what these points might be, you would have tried to include them in the original brief. Therefore they should be unexpected. Try to hold your study team together until after the presentation, so that you can quickly swing them into action to collect or analyse data relevant to these additional questions, and provide the answers before management have forgotten the original presentation.

If it sounds as though we had all these ideas worked out in advance, it is not so. We groped our way into some, we stumbled over others, and we only understood certain ones by analysing them after the event. But if by presenting them now we have saved you some of our wasted efforts, our experience will have paid for itself.

PAPER 2

EXPERIENCE WITH A PILOT SCHEME AND THE TRANSITION STAGE TO FULL MECHANISATION

by

S. C. Schuler

Head, TIL Reports Centre, Productivity Services & Information Division,
Ministry of Technology, Orpington, UK

SUMMARY

Discusses changes in operations and some of the problems occurring in the transition from purely manual operation, through partial mechanisation, to a proposed comprehensive computer system in an organisation handling 46,000 new titles annually. Partial mechanisation has involved the introduction of tape typewriters for typing bibliographic data and the production of printed indexes from these data using a bureau computer. The comprehensive system requiring an "in house" computer will permit SDI, subject searching, announcement bulletin and index production, and stock control to be carried out.

SOMMAIRE

L'auteur décrit les modifications de fonctionnement et les problèmes survenus au Centre des Rapports du TIL, au cours de la phase de transition entre opérations manuelles et un projet de système complet d'ordinateur, en passant par une mécanisation partielle. Ce Centre est une organisation assez importante traitant des rapports de recherche et de développement soit librement accessibles au public soit classifiés; le stock de rapports ainsi traites s'élève à presque un demi-million par an, et augmente annuellement de 46.000.

C'est par le système manuel que s'effectuent l'acquisition et la distribution des rapports, et l'annonce de leur parution dans un bulletin bi-mensuel ainsi que dans les listes spécialisées; c'est également manuellement que fonctionne un service d'informations techniques fournissant des listes de références en réponse à des demandes spécifiques, et des bibliographies de rapports traitant de sujets d'intérêt général et actuel.

Le premier stade de la mécanisation a impliqué l'introduction, dans le cycle de traitement des rapports, de quatre perforatrices sur bandes utilisées pour inscrire les principales données bibliographiques de chaque rapport dans la préparation des bulletins d'annonce. La bande perforée ainsi produite est traitée par un bureau d'ordinateurs en vue de l'etablissement d'index imprimés. L'etablissement de ces index a permis de répondre à un besoin des usagers et de réduire considérablement le classement des fiches dans les divers index de références.

Le stade final de mécanisation consistera en l'installation d'un ordinateur au sein du Centre. Cette installation a déjà été autorisée. Le système complet permettra d'effectuer toute une gamme d'opérations: diffusion sélective de l'information, recherches par sujet, préparation des bulletins signalétiques et des index imprimés, et contrôle des stocks.

EXPERIENCE WITH A PILOT SCHEME AND THE TRANSITION STAGE TO FULL MECHANISATION

S. C. Schuler

1. INTRODUCTION

During the past four years, pilot studies have been made on the benefits to be gained from using a comprehensive computer system for document handling and retrieval procedures at TIL Reports Centre of the Ministry of Technology (Mintech). These studies have led to the introduction of some first phase mechanisation and the use of a computer bureau for the production of printed indexes to the report collection. The decision to go ahead with the procurement of a computer was recently made and this paper describes the pilot schemes and the transition stage now reached prior to full mechanisation.

2. FUNCTIONS OF TIL REPORTS CENTRE

The Centre is Mintech's clearing house for unpublished report literature and its main functions are:-

(i) To exploit the large collections of unpublished reports on Science and Technology, both British and overseas.

(ii) To meet the technical information needs of Mintech, of other Government Departments (particularly the Ministry of Defence), and their Contractors, and other sectors of British industry.

(iii) To operate technical information exchange agreements with the USA, the Commonwealth and NATO countries.

The organisation of the Centre is shown in Figure 1 and this includes some statistics relating to its overall operations. It will be seen that the stock of reports is close on a half million titles and is growing at the rate of 20,000 per year in hard copy and slightly more for reports on microfiche. The reports are available on loan to users or for retention for the cost of reproduction.

The aim of the Centre is to collect all reports likely to be of interest to its varied range of customers and to make them as widely available as possible. This is done through three basic services; regular publication of abstracts and lists of material collected, distribution of a series of leaflets on specific items of technical innovation under the title Techlink, and a technical inquiry service which includes the compiling of bibliographies to customers' requirements.

The major publication of the Reports Centre is "R & D Abstracts". This is a twice-monthly accessions journal containing abstracts of the reports received at the Centre. The subject coverage of the journal is classified into 22 COSATI subject fields. It is prepared in various editions and an unlimited edition, listing on average 300 openly available reports, is on sale.

The reports announcement service is complemented by the Techlink service which provides a rapid and selective communication of information on new ideas, equipment processes and materials to individuals in British industry. A one or two page leaflet "Techlink" (an example is shown in Figure 2) contains the essentials of a particular technical innovation. The material is gathered from the Centre's collection of reports, from the United Kingdom Atomic Energy Authority, the National Research Development Corporation and from the Ministry's own research and development establishments. Each leaflet gives details of where and how further information can be obtained.

Techlinks are designed as a personal service restricted to those in British industry who are concerned with research, design and development and who can directly exploit the information presented. The subjects covered are classified under 52 headings and customers receive only those Techlinks that are relevant to their interests.

The technical inquiry service is based on a comprehensive card index to the report holdings maintained at the Centre. Subject searches are made on request and bibliographies in several subject fields are published at regular intervals.

3. LIMITATIONS OF MANUAL SYSTEM AND NEED FOR MECHANISATION

The manual system involved the maintenance of five separate card catalogues. (Title, Personal Author, Originator, Contract Number and UDC Subject). The cards (now over one million) were produced from the plates used for printing the "R & D Abstracts" journal and approximately thirteen thousand were supplied fortnightly.

The manual sorting and interfiling of new cards was a formidable task with no prospect of increase in clerical staff. Other limitations of the manual system were a lack of printed indexes to "R & D Abstracts", a pressing need for larger numbers of subject searches, inadequate bibliography service and shortcomings in the stock control and management statistics.

It was evident that many of these problems could be solved by automation which would also offer potential for fuller services. Accordingly, a small team was set up in 1965 to report on the feasibility of applying automatic data processing within the Centre. Several reports have been produced including TIL Report No. 12[1], describing the input and output problems; Report No. 16[2] and Report No. 18[3] which describe the main findings of the feasibility team.

In 1967 it was decided to go ahead with Phase 1 of mechanisation which involved the introduction of tape typewriters into the report processing cycle and the processing of the punched paper tape at a computer bureau for the production of printed indexes. During the period of Phase 1 financial approval was being obtained for installation of an in-house comprehensive computer system, the main services of which are shown in Figure 3.

4. IMPROVED PROCEDURES (PHASE 1)

The first stage of Phase 1 was to obtain tape typewriters for typing the main bibliographical data. The individual (first generation) tapes produced would be then edited and re-run through the tape typewriter to produce offset litho plates for printing the "R & D Abstracts" bulletin. The system adopted separated the various data fields in each bibliographical record by an Item Separator Symbol (ISS) (an inverted delta V). Concurrently a continuous second generation tape containing bibliographical details only would be punched for input to the computer bureau which would sort the data into the order required for printing the various indexes to "R & D Abstracts". Examples are shown in para. (q) of Annex 1.

4.1 Tape Typewriter System

After a survey of machines already in production, an order was placed for four Vonamatic tape typewriters from Business Data Products Ltd., and delivered during the later part of 1967. Figure 4 shows one of the machines in use. Operator training and tape program development occupied a period of some months, but the system became fully operational by June, 1968.

The Vonamatic tape typewriter utilises an IBM Model "C" standard electric typewriter. The models supplied to TIL each have one punch and two readers capable of reading either punched tape or edge-punched cards. ISO-7 eight-channel code is used. In the first four machines functional and programming keys are arranged in a panel on the working surface to the right of the typewriter and these include a coding key which when depressed allows functional codes to be punched into a tape. Operating speeds are:- typing: 600 characters per minute; reproduction:- 900 characters per minute; tape skip:- 1,200 characters per minute. One reader is housed in a well at the left of the typewriter and the second reader is a separate unit which can be moved to any convenient position on the working surface. The punch is at the right hand side of the control key panel. Override switches are fitted at the front of the desk to the right of the typist. These include skip and stop overrides, and a switch which will prevent the ISS symbol from being printed, although it still reproduces on a second generation tape which is being punched. A fifth machine was subsequently obtained. This is an improved model using solid state circuits and the reader and punch are in a separate unit to the right of the typewriter desk, the control keys are in a row in front of the reader and punch.

4.2 Problems Encountered with Machine Operations

Initially teething troubles were experienced with the tape typewriters. Intermittent failures of the punch units were caused by high room ambient temperatures and the close proximity of an electricity substation making the power supply voltage always on the high side. Voltage stabilisers were fitted and better room temperature control effected to give an improvement.

It cannot be too strongly emphasised that the computer tapes must be accurate. The tapes must be free from parity errors (i.e., an odd number of holes in one code), as such an error causes the tape reader of the computer to reject the tape. The computer program counts both the number of ISS symbols and also of line feeds (carriage returns) after certain fields in each item. To ensure that these are accurate, it is, therefore, necessary to read each computer tape through one of the typewriters and print out the contents. An occasional parity error can be corrected using a hand punch (if the error is due to a missing hole this is punched in to give the correct character; if there is one hole too many the character is deleted by punching all channels), but if this causes deletion of an ISS symbol the whole item must be 'killed'. The "kill item" code (Channels 2:3:6:8) is punched by hand in the tape feed about one inch before the first square bracket of the item in question. The computer will then ignore this item which can be fed in at the end of the next tape to be punched. It was hoped that after the first few batches of computer tape these checking runs could be dispensed with, but this has not proved possible. Reading a 50-item tape takes from 50-60 minutes, so that if the tape has to be repunched and re-read, the whole operation can take up to three hours.

4.3 Report Flow

The document flow processes for Recording, Preparation of "R & D Abstracts" and the Production of Printed Indexes are shown in Figure 5. The detailed steps in the operations are given fully in Annex 1.

During the implementation of the improved procedures, it was necessary to revise methods of recording and redesign some of the standard forms in use. Typical new routines include the following:

(a) *Movement Card*

To eliminate the step of writing the movement card, it was decided that typing the movement card and abstract would take place in one operation. A carbon-patched flimsy-movement card combination was designed and packs of these card-flimsy pairs obtained. The size of the carbon patch was such that when the bibliographical details and abstract were typed on the flimsy only the bibliographical details appeared on the movement card.

(b) *Process Sheet*

In order to reduce clerical work, it was decided to adopt the system of marking up the title page of a report so that as much information as possible could be taken by the typist direct from the report. The process sheet was redesigned to combine the functions of the original abstract and process sheets.

(c) *Edge-punched Originator Cards*

A standard entry for the many names and locations of originators of reports held by the Centre is now achieved by using edge-punched cards. Each originator is allotted a six figure code number. This number, selected from the TIL List of Corporate Authors relates to a punched card for that originator. A master set of cards suitably identified is held by the tape typists. The six figure code number is not typed on the carbon backed movement card but is transferred in non-print form to the first generation tape. When the first generation tape is used to produce the off-set litho plates for "R & D Abstracts" the six figure code number is again not typed but is transferred to the second generation tape.

The movement card, abstract and originators edge-punched card together with the tape are placed in a small wallet as shown in Figure 6 for subsequent vetting by the proof checkers and sorting into subject order by the bulletin editor.

4.4 **Preparation of "R & D Abstracts"**

"R & D Abstracts" bulletin is offset-litho printed from ... lith plates typed directly by the tape typewriters. The detailed steps in the operations are described in para. (p) of Annex 1. This method of direct typing on to printing plates was chosen as giving the best results with the reprographic equipment currently available at the Centre.

4.5 **Production of Printed Indexes**

(a) A contract was let to International Computers Ltd. (ICL) who prepared the computer programs necessary for the production of author, corporate author, report number, conference, translations and title indexes in quarterly and annual cumulations, from the data punched on the paper tape. These form the main clerical parts of the indexing system; the provision of detailed subject indexes posed other problems which are dealt with in (b) below.

Arrangements were made with the Central Computer Bureau of H.M. Stationery Office, for the provision of computer time. Each fortnight, paper tape containing the bibliographical data for an issue of "R & D Abstracts" is sent to the Bureau. At the end of each quarter, i.e., 6 issues of the bulletin, the Bureau provides the print-out for the six indexes. At the end of each year, data from the first three quarters and the final quarter are cumulated into annual indexes. The computer print-out is sent to an agency for the production of the bound printed indexes. This involves a 56% reduction of a double page of print-out on to a foolscap plate, printing, collating and binding. The early trial runs, although usable were very disappointing in their legibility. The ordering of entries was quite satisfactory and the poor legibility stemmed from a combination of factors:-

(i) The type fount on the printer used at the Bureau was not ideal for the degree of reduction required for printing.

(ii) The printer was run at its fast speed of 1,200 lines per minute which resulted in letters jumping out of alignment and other irregularities.

(iii) The print ribbon slipped resulting in the loss of the initial letters of lines.

(iv) The typeface was not always clean causing letters such as B, P, E to become filled in.

Attempts to offset the poor quality of the print-out by taking extra care in production of printing plates resulted in a slowing down of the printing operation.

In a subsequent run which tested the annual cumulation program on data for a half-year, steps were taken to improve the quality of the print-out. In particular, the printer was run at the slower speed of 800 lines per minute, and as far as possible the run was made immediately after the computer printer had been serviced.

A much better print-out resulted which gave very satisfactory indexes.

(b) A specification for a printed subject index based on descriptors has been drafted and a programming contract placed with ICL.

The first requirement in developing this index was an authoritative thesaurus of descriptors. The proposed LEX Thesaurus seemed to offer this and with the advantage that much of TIL's input from the USA would eventually be indexed with LEX terms.

The LEX Thesaurus was received in December 1968 and descriptor allocation commenced

The subject index will contain subject headings formed from single descriptors or up to three descriptors in combination. Skilful editing is required to give consistent headings.

In the computer input, the start of an index heading in the list of descriptors is indicated by an asterisk and the end of a heading by an oblique stroke. If two or three descriptors are used in the heading, they are separated by commas. For information retrieval as distinct from subject index compilation, commas and obliques can be given equal value as descriptor separators.

In this example of input the underlined descriptors are the index headings which will be listed with up to 82 characters of report title, the TIL Accession number and "R & D Abstracts" location number in the final subject index.

> /* <u>Ocean waves, mathematical models</u> /* <u>Shallow water, Hydrodynamics</u>/ Two-dimensional flow/ Beaches/ Deformation/ Numerical methods and procedures/ Programming (Computers)/ Mapping V.

4.6 Techlink Service - Use of High Speed Addressing Machine

Although not strictly part of the Phase 1 plan, an Elliott Dymatic 7000 high speed addressing machine has been installed to cope with distribution to the 4500 Techlink recipients.

Addresses are printed from small stencil type address cards. Outside the typing area of the card is space for punching up to 60 holes to provide coded information for selection of addresses.

For Techlink operation the holes are used as a series from 1 to 60 and each hole represents a specific subject code desired by the addressee. The holes can be programmed to convey other information.

The machine can address postcards, envelopes or print lists on roll paper at a speed of 7000 addresses per hour.

5. TRANSITION TO COMPUTER OPERATIONS (Phase 2)

Having now gained experience in the Phase 1 stage of mechanisation work has been proceeding on the preparation of an operational requirement for the computer and the necessary case writing to obtain financial support for the project to proceed.

5.1 Operational Requirement

A typical computer configuration to carry out the Centre's operations would consist of a central processor having a 64K character core store with two tape decks and two, or possibly three, exchangeable disk stores, depending upon the way in which the files of data are to be held and manipulated. A punched paper tape reader will be required and also, probably, a card punch, and output will be via a fast line printer capable of printing in upper and lower case and offering about 120 print positions. The possibility of carrying out the main printing operations in an off-line mode is being investigated and means are being sought for providing output suitable for driving computer-controlled photo type-setting equipment at some future date.

The main elements of the input program are shown in the Operational Requirement, (Annex 2) and programs for setting up files, sorting and printing data, retrieving information and carrying out control operations will be specified when the choice of equipment has been made. The extent ' which "program packages" can be made available by manufacturers will inevitably influence this choice.

5.2 Planning Networks

Defining the operational requirement is an important stage but equally important is the way in which the computer, when ordered, is integrated into the organisation.

It is necessary to define the problem and to plan the whole project so that each task involved in implementing the project is performed at the correct time and in a prescribed manner.

There are several key stages involved in the drawing up of plans and these include:

(i) Group the individual tasks under functions and areas of responsibility to give a better understanding of the work and define a network of the plan.

(ii) Agree on the staff resources necessary and estimate the duration of each activity and calculate the critical path.

(iii) Define the job specification and responsibilities for each staff member required to implement the plan. Set up reporting control and progressing procedures.

(iv) Prepare plans and charts to utilise staff, establish separate charts to progress each function and relate these to the whole project.

The Pert Chart (Jan. 1969) for the early procurement stages and the recruitment and training of staff for the TIL Computer is shown in Figure 7.

5.3 Looking Ahead

The pilot stage of mechanisation at the TIL Reports Centre is now merging into the computer procurement stage. Discussions are currently taking place with computer manufacturers, four of whom have made preliminary submissions. These discussions are designed to relate recent developments in information retrieval techniques to the Centre's needs and the final specification will be drawn up having regard to the extent to which proved hardware and ready made programs can be incorporated.

An illustration of the Centre's current overall processes and services is given in Figure 8, the extra services to be provided when the on-site computer is fully operational are indicated in the 'close lined' areas.

As a further longer term development the Centre's computer could be used as Mintech's focal point for interconnection with other Information Centres on the lines shown in Figure 9.

ACKNOWLEDGEMENT

The author would like to acknowledge the help and suggestions in the preparation of this paper of Miss E.H. Ridler, Mr V.J.Benning and Mr R.H.Howe of TIL Reports Centre.

REFERENCES

1. Isotta, N.E.C. *Mechanisation and Documentation Input/Output Problems.* TIL Report 12, Oct. 1964, 49pp.

2. Rogers, V.J. et al. *Automatic Data Processing Study.* TIL Report 16, April 1966, 38pp.

3. Rogers, V.J. et al. *Computer Control of Documentation - Feasibility Study Report.* TIL Report 18, May 1967, 16pp.

4. Ridler, E.H. et al. *Production of R & D Abstracts and Printed Indexes using Tape Typewriters.* TIL Report 21, Aug. 1969.

ANNEX 1

DETAILED PROCEDURE FOR TAPE TYPEWRITER SYSTEM

The main steps in procedure now in operation are:

(a) After duplicate check, a new report is given an accession number. The number is stamped on a movement card. A temporary 5 in. × 3 in. card for the Originator card catalogue is written.

(b) The report title page is annotated according to the process sheet layout. Items to be recorded are ringed and numbered according to the numbered fields in the process sheet. The relevant boxes in the movement card are ticked when the typist has an entry to record.

(c) Edge-punched originator card is selected.

(d) Carbon-backed flimsy + movement card + originator card are placed inside a wallet and together with the process sheet are placed inside the report.

(e) Report passed to subject analyst.

(f) Analyst either annotates an author abstract in the report or writes an original abstract on the process sheet.

(g) Analyst allocates UDC numbers and Descriptors using the LEX Thesaurus and writes them on the process sheet.

(h) Abstractor writes his recommended distribution on the movement card.

(i) Report + process sheet + wallet containing the cards is passed to the tape typist.

(j) The typist has the tape program permanently in Reader 1 of her machine. She removes the flimsy and movement card from the wallet and places it in the typewriter. She places the edge-punched originator card in Reader 2.

(k) The typist carries out a prescribed series of operations; typing the items of bibliographical information from the title page of the report and calling in the tape program by pressing the "Start Read" button of the machine after completing each item of data.

(l) After the abstract is completed and the tape program run to the end, a short length of tape feed is punched, the tape torn off, folded and inserted in the wallet together with the flimsy + movement card and the originator card. The wallet and process sheet are returned to the report which is passed to the checkers.

(m) The checkers mark small corrections on the flimsy. For major errors, the report is returned to the typist to punch a new tape.

(n) The checked abstracts plus reports are passed to Editor, "R & D Abstracts" where the flimsies are detached and the movement cards returned with the reports to the Distribution Section. The flimsies are kept in the wallets with their respective tapes.

(o) Editor, "R & D Abstracts" edits abstracts, verifies UDC numbers and descriptors and allocates appropriate COSATI field and group number. Abstracts are sorted into field and group order and numbered serially with an item number.

(p) The wallets containing the flimsies plus tapes in their serial order are passed back to a tape typist. The first tape is placed in the reader and an offset litho plate placed in the machine. The first subject heading is typed from the keyboard with the punch switched off. The first abstract is then run through the reader, the punch automatically being switched on by a code at the start of the tape inserted originally from the tape program. A code punched in the tape from the program automatically stops the tape after the first square bracket so that the item number can be inserted from the keyboard. At the same time as the plate is typed, a second generation tape is punched but a code in the report tape inserted by the tape program automatically switches the punch off before the abstract so that only the bibliographical details appear on the computer tape. The punch is automatically switched on at the end of the abstract and the machine stops so that the typist can press the Tape Feed button to give a short length of blank tape before the next item. The next abstract tape is then inserted in the reader and the process repeated. Further subject headings are typed on the plates (but not on the second generation tape) as required. Each second generation computer tape normally contains fifty items, an arbitrary number chosen for convenience in handling. At the end of each tape the typist types TC4, a signal to the computer that the end of the tape has been reached. Errors in the original report tapes indicated by the checkers and editors can be corrected during typing the plates by switching the reader into single step before the error, typing the correct characters manually and moving the tape the correct number of codes past the error.

(q) **Examples of Tape Typewriter Output**

(1) *First Typing*

```
[]   ERDE-13-R-68  ▽   ▽              ▽ UNLIMITED
▽      Explosives Res.+Dev.Est., Ministry of Technology, UK
▽    COMPATIBILITY OF WHISKERS OF SILICON NITRIDE, SILICON CARBIDE AND ALUMINA WITH METALLIC
     MATRICES AT HIGH TEMPERATURES
▽
▽
▽      Cannell, J.C.        8.1968     50pp 15ref
▽      ▽    ▽ WAC-189-011
▽           ▽
▽
▽    UDC.  669.018.95/661.665.1/661.862.22/546.284'171
▽    This report describes the continuation of a series of high temperature compatibility trials,
     screening various combinations of whiskers and metals for prolonged service at 1100 °C, as
     a preliminary to possible development of composite turbine blades. The trials showed that
     only one sample of alumina whiskers remained stable even in pure nickel.  GMH
```

(2) *The tape from (1) is run through the machine cutting offset litho plate and producing 2nd generation tape*

The following is the R+D Abstracts entry:

```
[1001-6901]  ERDE-13-R-68                     UNLIMITED
       Explosives Res.+Dev.Est., Ministry of Technology, UK
     COMPATIBILITY OF WHISKERS OF SILICON NITRIDE, SILICON CARBIDE AND ALUMINA WITH METALLIC
     MATRICES AT HIGH TEMPERATURES

       Cannell, J.C.        8.1968     50pp 15ref
             WAC-189-011
```

UDC. 669.018.95/661.665.1/661.862.22/546.284'171
This report describes the continuation of a series of high temperature compatibility trials, screening various combinations of whiskers and metals for prolonged service at 1100°C, as a preliminary to possible development of composite turbine blades. The trials showed that only one sample of alumina whiskers remained stable even in pure nickel. GMH

(3) *Print-out of 2nd generation tape*

[1001-6901] ERDE-13-R-68 ▽ ▽ ▽ UNLIMITED
▽ 445000Explosives Res.+Dev.Est., Ministry of Technology, UK
▽ COMPATIBILITY OF WHISKERS OF SILICON NITRIDE, SILICON CARBIDE AND ALUMINA WITH METALLIC MATRICES AT HIGH TEMPERATURES
▽
▽
▽ Cannell, J.C. ▽ ▽ 8.1968 ▽ 50pp 15ref
▽ ▽ ▽ WAC-189-011
▽ ▽
▽
▽

ANNEX 2

OPERATIONAL REQUIREMENTS FOR A COMPUTER IN TIL REPORTS CENTRE

1. *The Task*

It is intended that the following work will be undertaken by the computer:-

(a) *Selective Dissemination of Information.* This service will be based upon the matching by computer of index terms defining a document's contents against similar index terms defining a user's field of interest. A successful match will result in a print-out of bibliographic information and details of the user's organisation and address.

(b) *Subject Searches.* Searches are carried out in response to requests which can be translated into computer index entries.

(c) *Special Bibliographies.* It is a necessary function of TIL to offer to its customers, from time to time, bibliographies concerned with particular subjects, authors, etc.

(d) *Printed Indexes.* Experience has been gained in the production of printed indexes by computer bureau. This operation makes use of punched paper tape produced by TIL tape typewriters.

(e) *The Printing of Abstract Bulletins.* "R & D Abstracts" is currently being reproduced from offset-litho masters created by the tape typewriters, but it is intended that the computer print-out will be used to produce the litho masters. Print quantity and format considerations are such that a single case fount is unacceptable, so that this task depends upon the provision of facilities for printing in upper and lower case.

(f) *Stock Control.* New documents will be subject to stock control by computer. Retrospection offers difficulties, but the processes of document contro[illegible] [illegible] at the records of existing reports for which there is a continuing demand may be gradually [illegible].

The above tasks will not necessarily be undertaken in the order shown.

2. *Activity*

A summary of the work load is given in Appendix 1 - File Contents. Other Appendices covering Input and Output Volumes are available.

Since the processing routines will depend upon the system and equipment employed, no attempt has been made to estimate the activity due to intermediate processing.

3. *Capacity*

The equipment must have the capacity to process within a single week's shift the average load imposed by the task defined in Sections 3 & 4 including at least 8 hours for parts (a), (b) and (c) of Section 3, and an allowance of 100% of the job time for parts (d), (e) and (f) to cover contingencies.

The configuration should be capable of considerable expansion.

4. *Equipment*

Computer equipment approximating to the following configuration is envisaged:-

(a) Central Processor having not less than 60K character core store (or its equivalent).

(b) A console typewriter.

(c) Punched paper tape reader capable of reading ISO7 code.

(d) Punched paper tape punch.

(e) Line printer:- at least 120 print positions; upper and lower case character set with special symbols.

(f) Magnetic tape and/or direct access magnetic storage capable of handling the input and output volumes indicated in Appendix 1 assembled into appropriate files. This equipment should include at least two tape decks capable of reading and writing 7-track magnetic tape at a packing density of 200 rows per inch, in accordance with BSS 3958/1966; capability of reading and writing at 556 rows per inch is also desirable.

(g) Voltage regulation may be required.

The tape typewriters used to prepare computer input produce tape in the ISO 7 code having even parity in the eighth channel.

The Department will consider the possibility of obtaining printing facilities separately from the main computer configuration.

5. *Equipment Trials*

Acceptance of the equipment will be subject to the satisfactory completion of equipment trials conducted by the Technical Support Unit of the Ministry of Technology on behalf of H.M. Treasury and H.M.S.O.

6. *Demonstration*

As part of the process of evaluation the equipment manufacturers may be asked to demonstrate selected tasks on the proposed equipment.

7. *Software and Systems Support*

Manufacturers will be asked to provide details of the systems and software support which will be made available.

8. *Program Testing*

Manufacturers must be able to offer facilities for testing programs prior to the delivery of the equipment, on a configuration similar to that ordered.

9. *Cost of Equipment*

The computer and peripheral equipment may be purchased or hired. If purchased the price is not expected to exceed £...,000. If it is decided to obtain printing facilities separate from the main computer configuration (see Section 4) the cost of printing equipment will be deducted from this sum.

10. *Delivery*

The date by which delivery of equipment can be offered by manufacturers will be an important factor in the evaluation of tenders.

11. *Location*

A computer is to be installed in Block i. of Station Square House, St. Mary Cray, Kent, in accommodation which has been set aside on the ground floor.

APPENDIX 1

FILE CONTENTS

The file structure must be compatible with the equipment employed. The following structure would enable the essential TIL operations to be carried out, but it is realised that alternative file structures could be effective.

1. *Processing of Main Bibliographic Information (MBI)*

Each document processed could constitute a record in a file held for this purpose. A maximum of 3,000 characters should be allowed for each record, the average being about 1,600 characters. The following data will be covered:-

	Max.	*Ave.*
Accession Number	20 Ch	14 Ch
Originator's Number	24	17
Other Reference Numbers (up to 3)	72	24
Originator's Identity (including country of origin)	72	48
Personal Author or Authors (up to 5)	130	52
Foreign Title	250	12
Title (English)	250	120
Conference Paper	250	25
Date of Report	8	4
Security and Availability Classification	24	2
Contract Number and Period	28	14
Project Number	24	12
Pagination	14	9
Abstract Location Reference	17	13
Descriptors	300	200
Abstract (see Note 1)	1500	1000
	2983 Ch	1566 Ch

Allowing 20,000 items per year for five years, the total current file length would be 100,000 items, i.e. a maximum of 300,000,000 characters approximately.

2. *Stock Control* (See Note 2)

Each record in the Stock Control File could be held in two parts as shown. The first part would identify the document and the second part would provide space for up to 40 transactions per document. The following data will be required:-

Document Identification	*Max.*	*Ave.*
Accession Number	20 Ch	14 Ch
Quantity of copies held	3	2
Date of receipt	8	8
Pagination	14	9
	45 Ch	33 Ch

Transaction		
Copy Number (or Quantity of Copies in Transaction)	4 Ch	1 Ch
Customer Identity	6	6
Loan Data	3	2
Date of Transaction	8	8
Subsequent Action Data (loan data and date)	11	6
	32 Ch	23 Ch

Allowing 20,000 items per year for five years, the total file length would be 100,000 items, i.e. a maximum of 5,000,000 characters of identification data and, say, an average of 30 × 2,500,000 characters of transactions. This represents a total of 80,000,000 characters approximately.

3. *Subject Searches*

Subject searches (information retrieval) will be based upon the use of descriptors, so that a compact file, relating descriptors to document reference numbers only, will be required for search purposes.

If the file is held in "inverted" form, then, allowing for a thesaurus of 15,000 descriptors having an average of 80 document reference numbers allocated to each in five years, the file will contain at least 5,000,000 characters.

4. *Printing of "R & D Abstracts" and Indexes*

The data required for the compilation of "R & D Abstracts" could be accumulated continuously and held in a separate file for such time as will be required for the printing and issue of the bulletin and its indexes.

The data required for the compilation of the Indexes to "R & D Abstracts" could also be held separately for the convenience of printing.

There will be four quarterly cumulations and an annual cumulation of the Indexes so that the file will require regular updating throughout the year. A fresh start will be made after the issue of the annual cumulation.

5. *Selective Dissemination of Information* (SDI)

Capacity will be required for storing customers' profiles (see Note 3) in the form of lists of descriptors. TIL serves some 7,000 customers, but it is expected that about 800 different profiles would be registered when the service is fully developed.

	Max.	*Ave.*
Descriptors per profile	300 Ch	200 Ch
No. of profiles	800	800
	240,000 Ch	160,000 Ch

Thus an initial space allocation of 200,000 Ch should suffice, with possible expansion to, say, 2,000,000 Ch as the system develops.

Notes

1. The abstract is not always required as part of the main bibliographic information. It may, therefore, be desirable to hold abstracts on a separate file with suitable cross references.

2. Stock Control transactions generally fall into two broad classes, i.e. current and completed. The stock control file could thus be held in two parts accordingly.

3. A "profile" is a definition of the field of interest of a customer in the form of a list of index terms (descriptors).

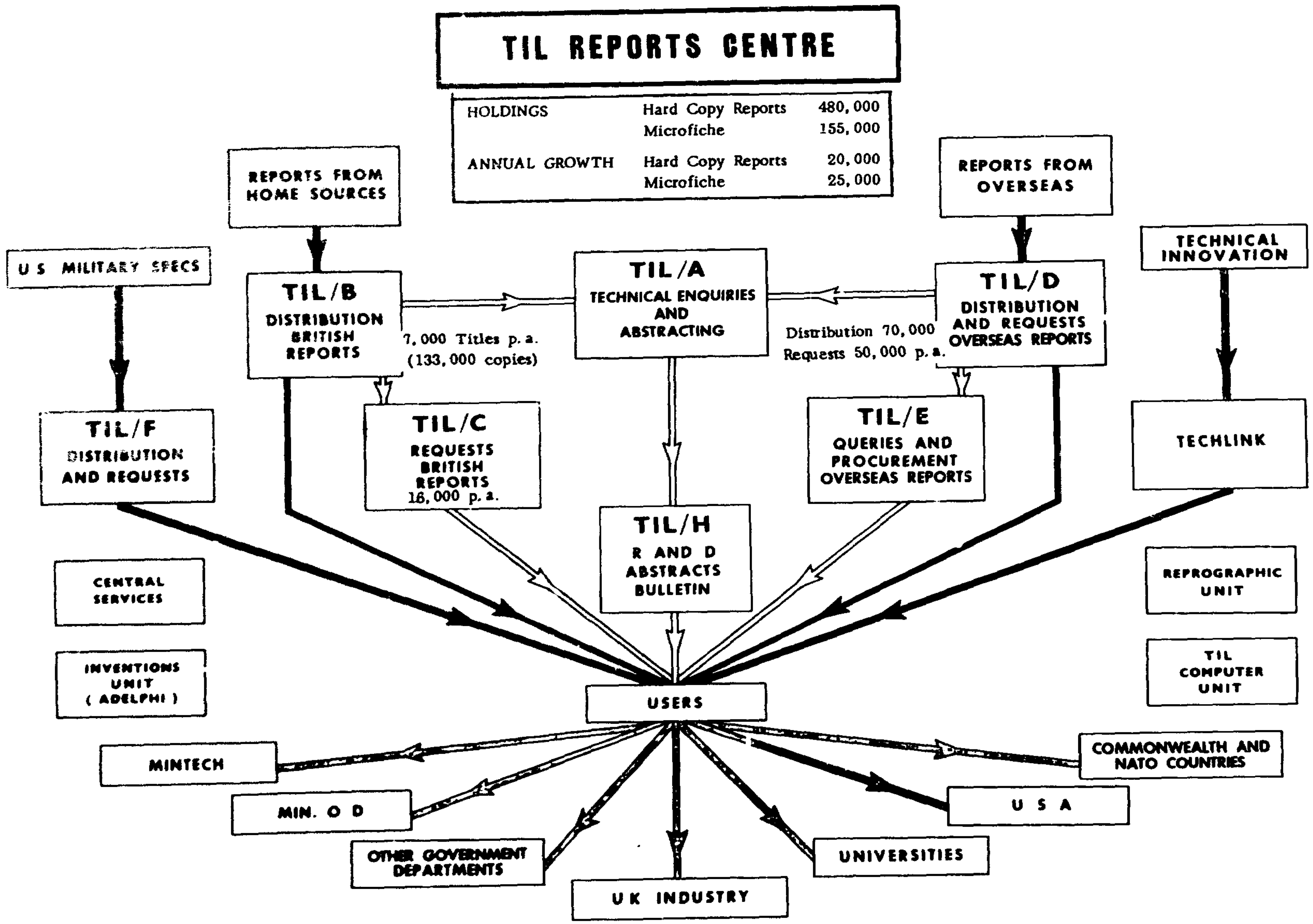


Fig. 1 TIL Reports Centre Organisation

research TECHLINK industry

MINISTRY OF TECHNOLOGY
Productivity Services and Information Division

Subject Code	9, 15
Serial No.	5

CONNEXION TO FLAT STRIP CABLE

This is a quick and simple method of making satisfactory connexions to the individual cores of multi-core flat strip cable. The connector combines the cores and insulation into a mechanically and electrically reliable unit on which control boxes and service outlets can be mounted.

The unit has a base on which the cable is located and tensioned by clamping strips. Over this is placed a connector plate having a series of toothed metal strips (the number depending on the circuits involved) and two transverse semi-cylindrical projecting bars corresponding to recesses in the base. As the connector plate is tightened down, the projecting teeth perforate the insulation of the cable and form wiping butt joints with the conductors as they are drawn inwards by the entry of the projecting bars into the mating grooves. The connexions so made are brought out to 'potted' multi-way terminals.

Since the heat dissipation of flat strip cable is superior to that of other types, voltages and loadings can be increased without disproportionate increases in size and weight. This method could therefore be incorporated in the modules of industrialized building systems to carry lighting, telephone and other circuits.

INFORMATION FROM

Signals Research and Development Establishment,
Christchurch, Hants.
Tel 0425 2 2361

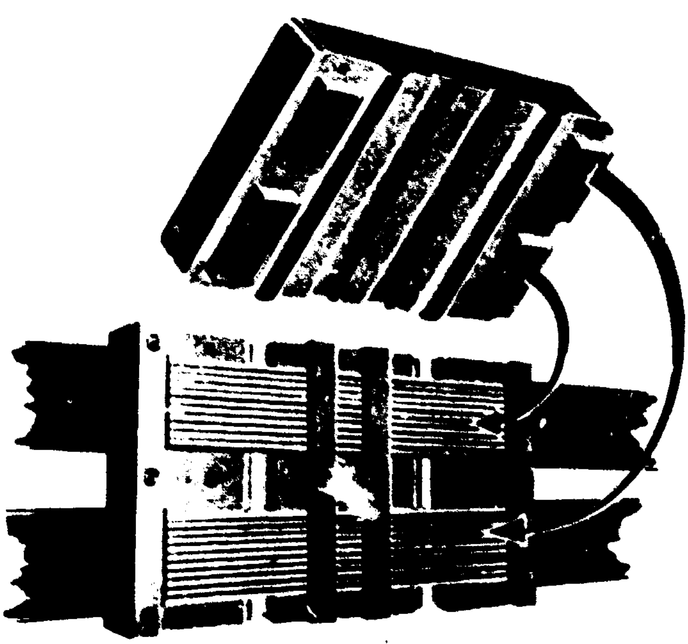


Fig. 2

Fig. 2 Example of a Techlink

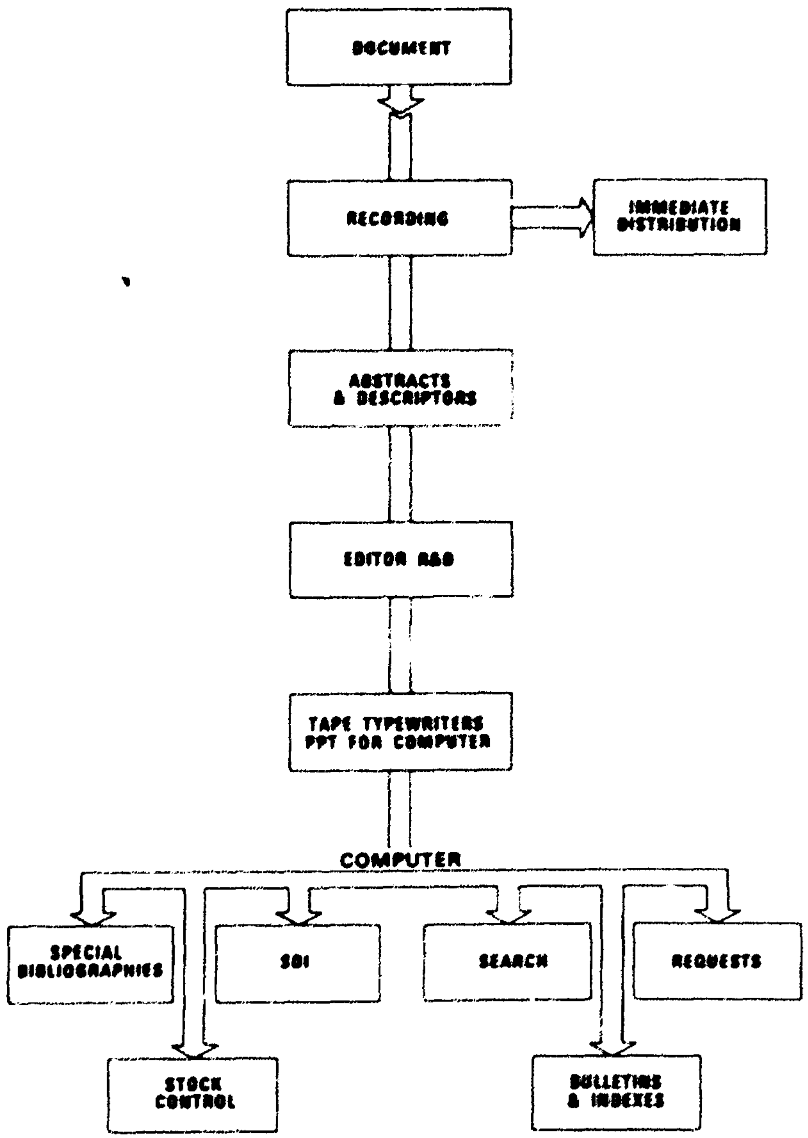


Fig. 3 Computer scheme flow diagram

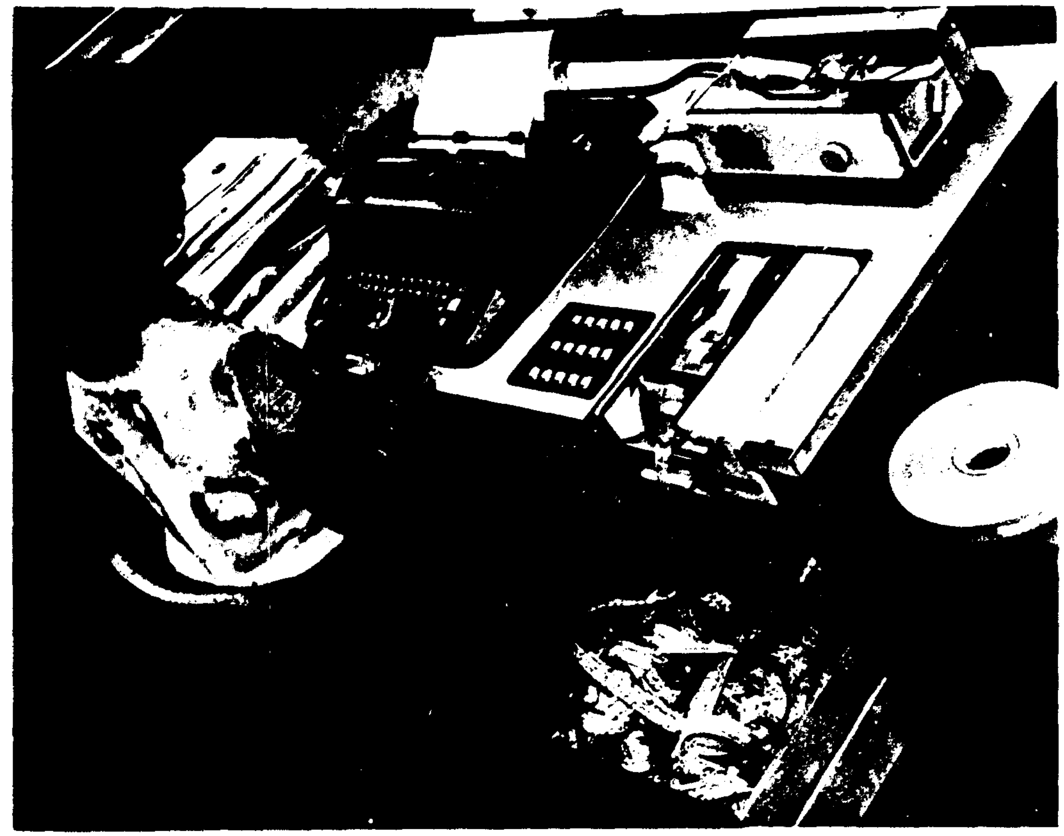


Fig. 4 Tape typewriter operation

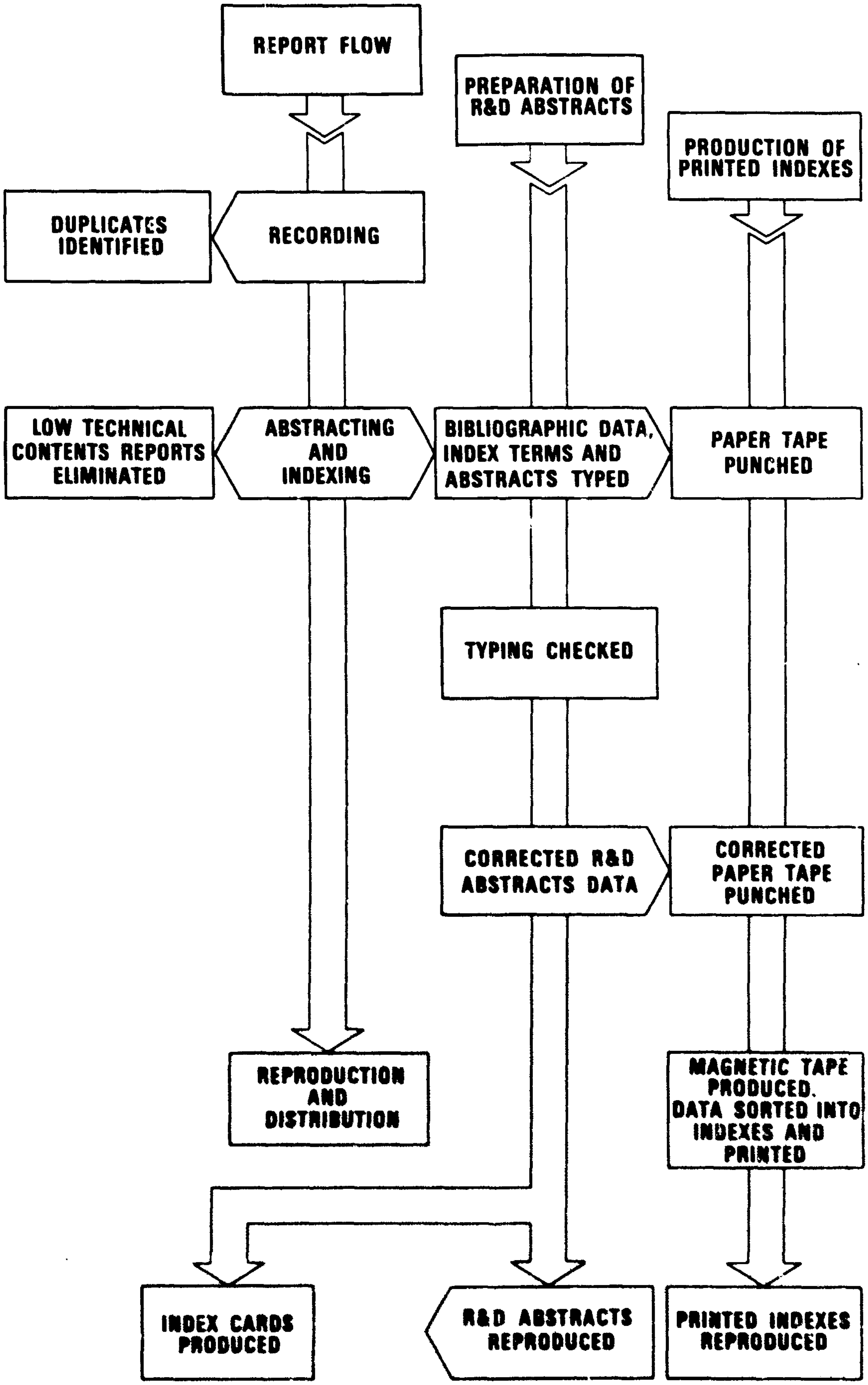
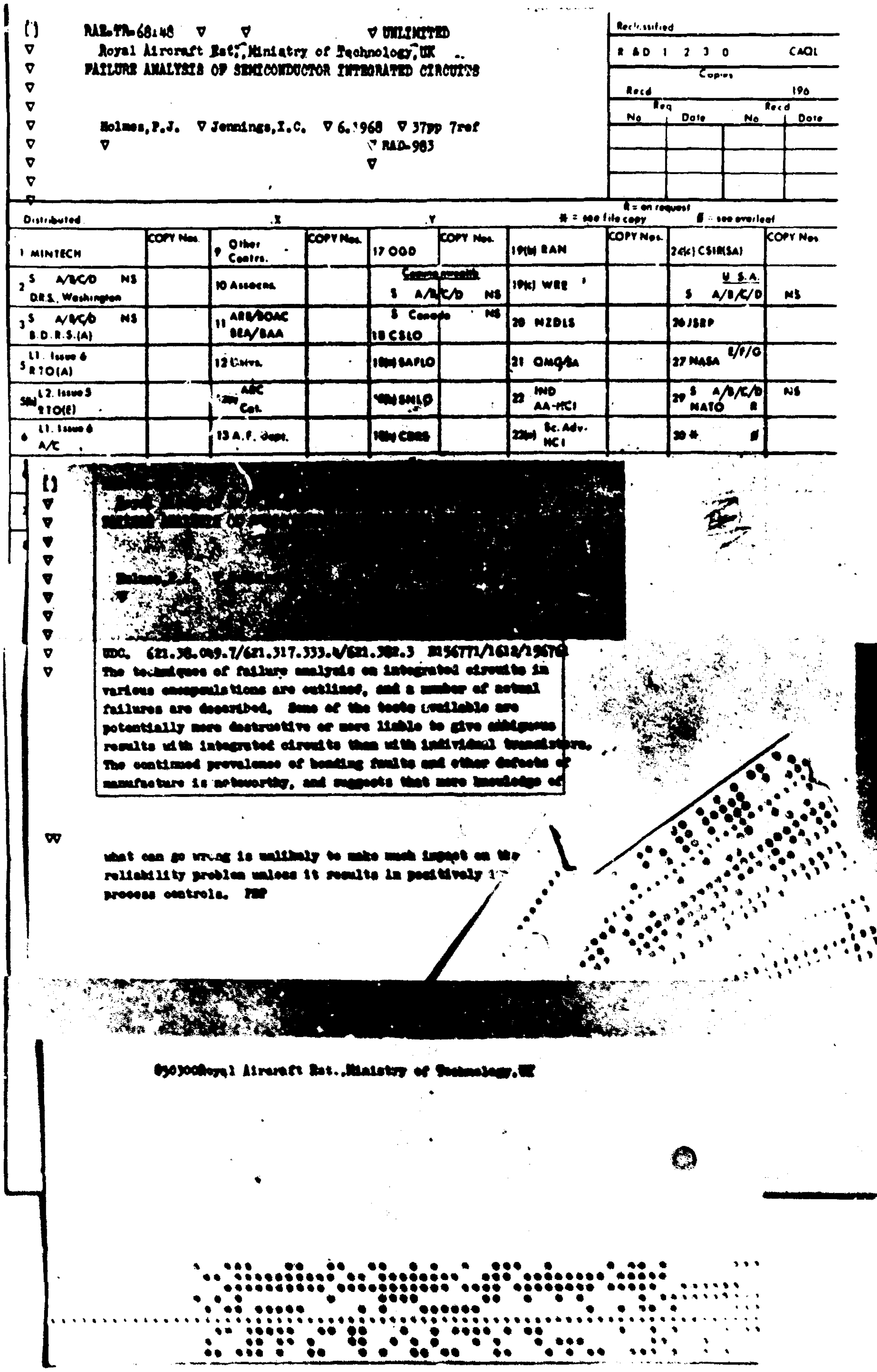


Fig. 5 Report flow processes



RAE-TR-68148 ▽ ▽ ▽ UNLIMITED
Royal Aircraft Est., Ministry of Technology, UK
FAILURE ANALYSIS OF SEMICONDUCTOR INTEGRATED CIRCUITS

Holmes, P.J. ▽ Jennings, I.C. ▽ 6.1968 ▽ 37pp 7ref
▽ RAD-983

Reclassified
R A D 1 2 3 0 CAOL
Copies
Recd 196

Req No	Req Date	Recd No	Recd Date

R = on request
Distributed X Y # = see file copy ≠ see overleaf

1 MINTECH	COPY Nos.	9 Other Contrs.	COPY Nos.	17 OGD	COPY Nos.	19(b) RAN	COPY Nos.	24(c) CSIR(SA)	COPY Nos.
2 S A/B/C/D NS D.R.S. Washington		10 Assocns.		Commonwealth S A/B/C/D NS		19(c) WRE		U.S.A. S A/B/C/D	NS
3 S A/B/C/D NS B.D.R.S.(A)		11 ARB/BOAC BEA/BAA		S Canada NS 18 CSLO		20 NZDLS		26 JSEP	
5 L1. Issue 6 RTO(A)		12 Univs.		18(a) SAFLO		21 OMG/SA		27 NASA E/F/G	
5(b) L2. Issue 5 RTO(E)		12(a) ARC Cat.		18(b) SNLO		22 IND AA-HCI		29 S A/B/C/D NATO R	NS
6 L1. Issue 6 A/C		13 A.F. Dept.		18(c) CDRS		22(a) Sc. Adv. HCI		30 #	≠

UDC. 621.38.049.7/621.317.333.4/621.382.3 RE56771/1612/19676

The techniques of failure analysis on integrated circuits in various encapsulations are outlined, and a number of actual failures are described. Some of the tests available are potentially more destructive or more liable to give ambiguous results with integrated circuits than with individual transistors. The continued prevalence of bonding faults and other defects of manufacture is noteworthy, and suggests that more knowledge of

what can go wrong is unlikely to make much impact on the reliability problem unless it results in positively improved process controls. PJH

050300Royal Aircraft Est., Ministry of Technology, UK

Fig. 6 Document movement card, abstract and corporate author punched card

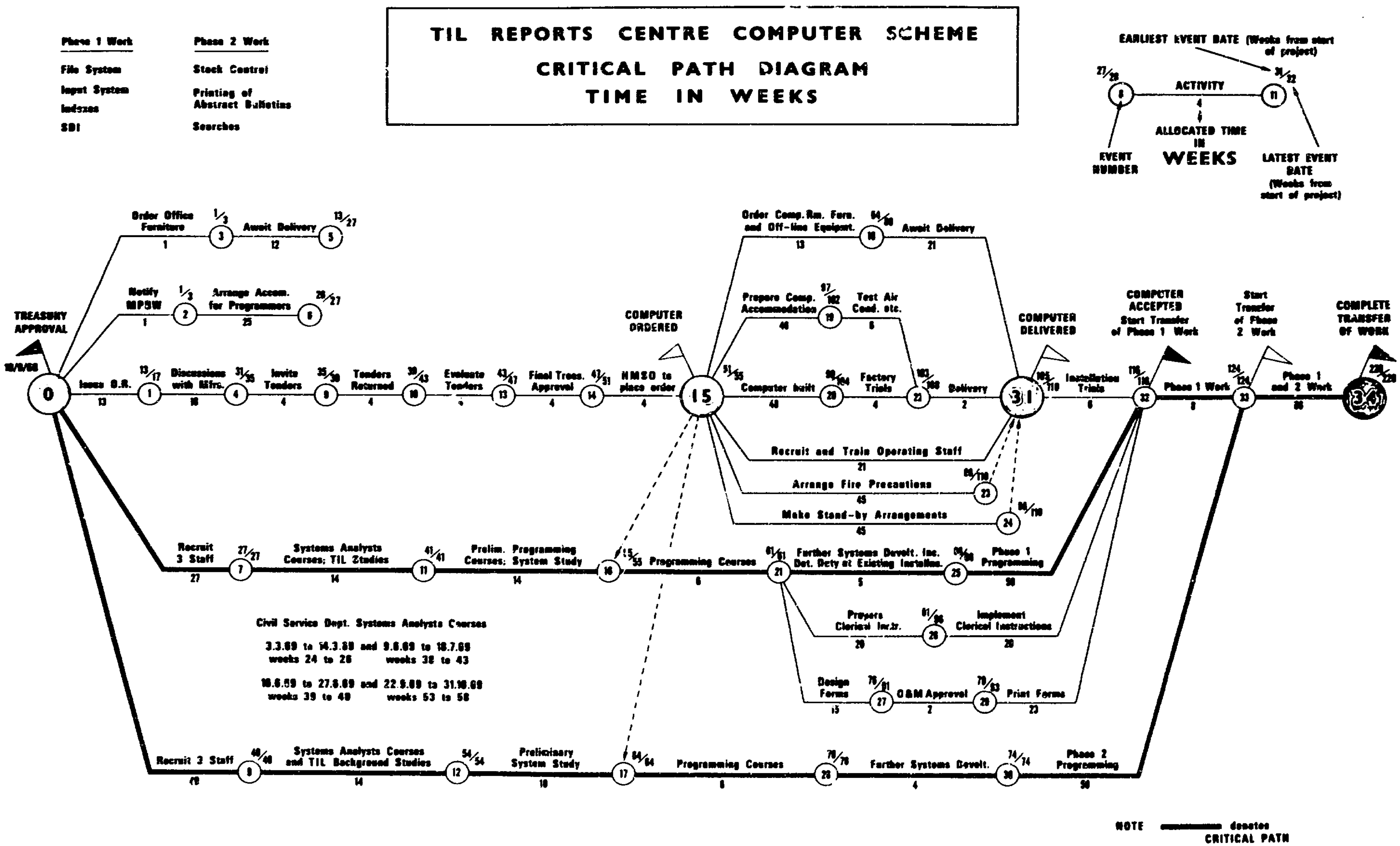


Fig.7 Critical path diagram

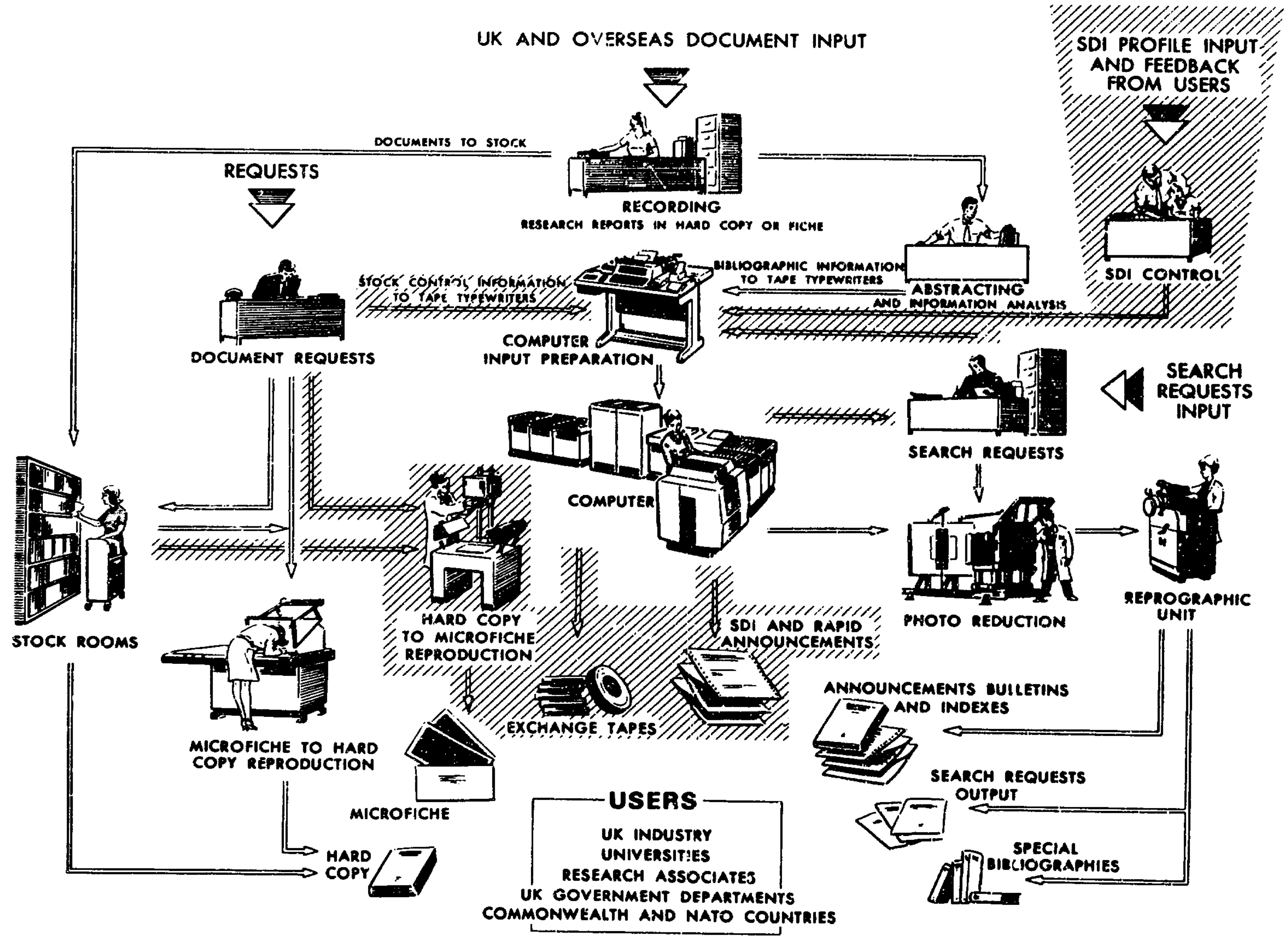


Fig. 8 TIL Reports Centre processes - current and projected

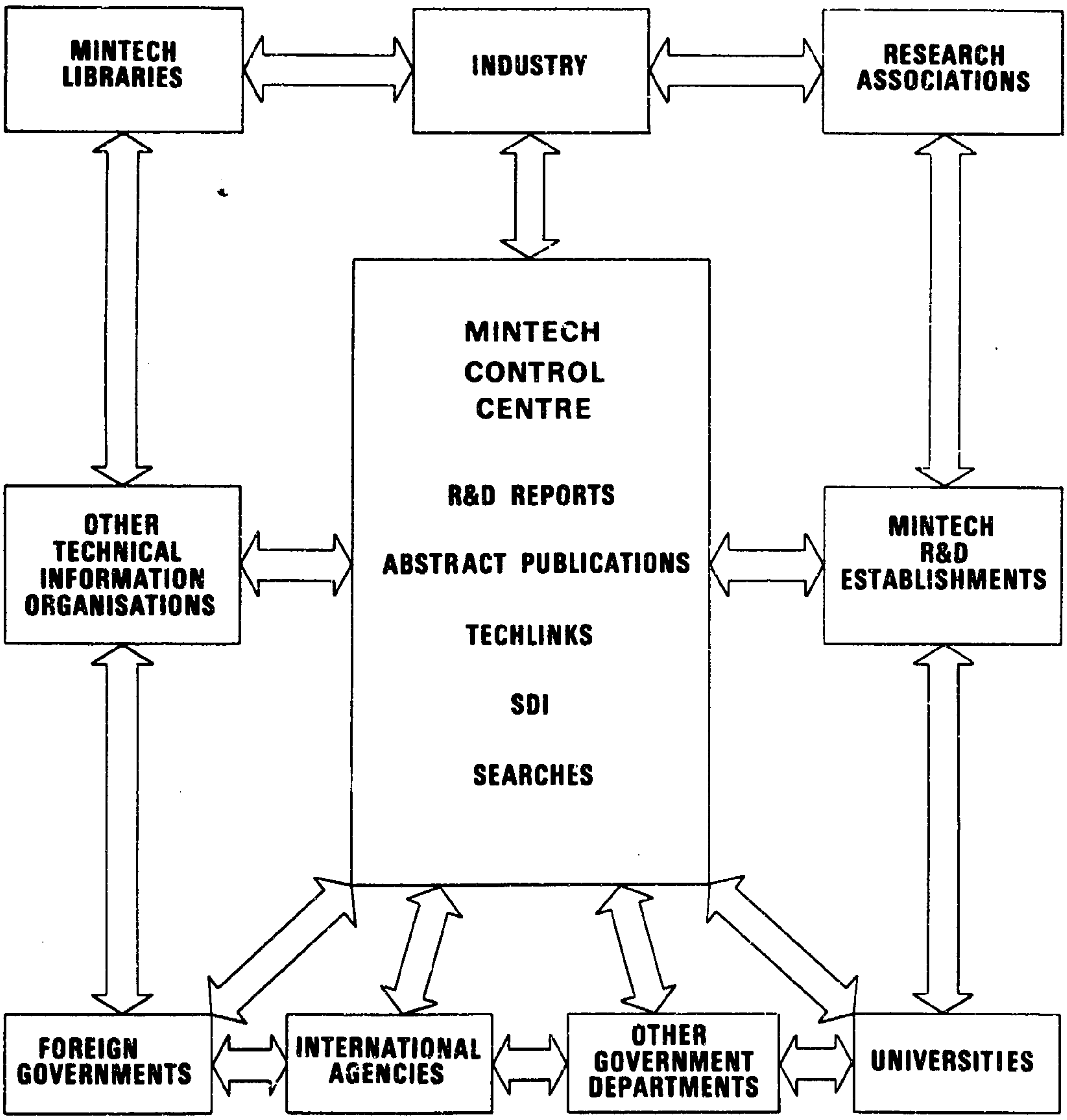


Fig. 9 Looking ahead - Mintech Data Control Centre

PAPER 3

THE USER AND THE AUTOMATED SYSTEM IN A EUROPEAN ENVIRONMENT

by

N.E.C. Isotta

Documentation and Library Service,
European Space Research Organisation, Neuilly, France

SUMMARY

The paper briefly outlines the early history and development of the ESRO/ELDO Space Documentation Service. The installation of the new direct access remote console system (NASA/RECON) is described. The necessity, and difficulty of talking to the multilingual European user is emphasised. Cost considerations are mentioned together with the user's requirements in this respect. The current feasibility and design study for the creation of a databank on electronic components and materials is outlined. The possibility of inter-disciplinary international direct access computer systems is suggested.

SOMMAIRE

Cet exposé retrace à grands traits les débuts et l'évolution du Service de Documentation Spatiale de l'ESRO/ELDO. L'auteur décrit l'installation du nouveau système à accès direct et pupître commandé à distance (NASA/RECON). Il souligne la nécessité de parler à l'utilisateur européen multilingue, et les difficultés impliquées. Il mentionne certaines considérations de coût, ainsi que les conditions requises par l'utilisateur dans ce domaine. Il expose brièvement l'étude de conception et de practicabilité actuellement effectuée et portant sur la création d'une banque d'informations sur les composants et matériels électroniques. Il suggère enfin la possibilité de systèmes d'ordinateurs inter-disciplinaires à accès direct.

THE USER AND THE AUTOMATED SYSTEM IN A EUROPEAN ENVIRONMENT

N.E.C. Isotta

INTRODUCTION

The European Space Research Organisation was established by the Convention of 14 June 1962, and came into force on 20 March 1964. The ESRO/ELDO Space Documentation Service was not immediately set up; it was first established in November 1965 so that the development period has been somewhat shorter. During 1968 ESRO launched three satellites and thus achieved for the first time, full operational status. Although modest by comparison with the NASA achievements, they represent the most complex satellites ever designed and built in Europe, or operated from Europe. Not surprisingly, due to its later initiation it is expected that the Space Documentation Service will only be providing all the services envisaged from 1970.

In a "Study of mechanization in DOD libraries and information centres" (Ref. 1) as long ago as September 1966 the following conclusion was drawn: "the survey data clearly indicate some cases in which a program of mechanization was initiated and pursued simply because it was the 'thing to do'. Several of the systems seemed remarkably able to operate and to expand in the vacuum of a nearly complete absence of considerations of user needs." It seems however, that in many of these cases what really happened was that machine time and programming effort were probably available free of charge and that therefore the 'thing to do', which was in any case interesting, became feasible, and even desirable in order to help justify equipment installed for other purposes. In the case of ESRO however, an existing computer facility was extensively upgraded in order to provide the additional power and flexibility which appeared to be necessary purely for documentation purposes. In this respect the ten ESRO member states as represented by the delegates to the controlling committees, were sufficiently far-sighted to see the real need for a system versatile enough to satisfy the varied information needs, on a centralised basis, of all the member states.

TO DESIGN OR NOT TO DESIGN

Before the establishment of SDS in November 1965, prolonged negotiations with NASA had taken place in order to arrange an exchange agreement between NASA and ESRO/ELDO. It was more than clear in the early history of the Organisation, that in the first place it would be impossible to build up quickly a useful reference file in the space field, apart from questions of cost, and that secondly the design of a new information retrieval system would take too long and would also cost too much. In the circumstances the most obvious and most logical arrangement was to make use of existing NASA facilities. As a result NASA generously entered into the information exchange agreement which in a strengthened and proved form still exists today. Under this agreement NASA provided the magnetic tape files of references to unclassified documents in the aerospace field (some 400,000 documents). In addition, master microfiche copies of the reports themselves together with the necessary computer programs to enable the files to be machine searched, were provided. In return ESRO, as far as possible, acts as a focal point for European aerospace reports. These are sent to NASA completely processed for direct input to the machine system.

The system developed by NASA was integrated into the European context with only very minor modifications. This was a linear file search system which operated as effectively as necessary on the then some 200,000 documents. For a new-born organisation such as ESRO this system in itself, which was operated in Paris under external contract, was of inestimable value. It has effectively satisfied the information needs of ESRO/ELDO staff and institutes in member states until July this year. However, even prior to this date it had become apparent that both the volume of searching required, and the size of the file itself warranted some fundamental re-design of the concept. Once again it was decided that ESRO/ELDO could not really afford the full effort necessary and we turned once more towards NASA. For some time NASA had been developing a large scale batch processing system for use on inverted files in an IBM 360 system. This it seemed would answer our requirements. However, an even more powerful development in the form of NASA/RECON made its appearance. With the concurrence of NASA sufficient funds were made available by ESRO to modify the existing NASA/RECON system, under contract with Lockheed, for operation in the rather different ESRO environment. The situation now is that the RECON system has been installed and is operational between Paris and the ESRO computer centre in Darmstadt, Germany. Flexibility of this new system is such, that it is immediately clear that the decision not to design or develop something else was correct.

TALKING TO THE USER

The recent report by The National Academies of Sciences and of Engineering on "Scientific and Technical Communication" (Ref. 2) states that "the management of all scientific and technical communication activities must be as responsive as possible to the needs, desires, and innovative ideas of the scientific and téchnical groups that they serve. These activities must be sufficiently flexible to adapt rapidly to changes in user needs and communication techniques." This is stated in the report, to be one of three main principles to guide the management and planning of scientific and technical information programmes and services. However, a vitally important point emerges in this connection, with regard to the difficulty of knowing the needs and desires etc., particularly in a multilingual European society. So long as the service concerned remains a purely passive one it is unlikely that the user will in general take the trouble to communicate his requirements to the organisation providing the service. A part of the trouble of course, lies in the fact that the user is normally not aware of changes or improvements that could be made, and consequently of the desirability of making them. For example 50 years ago the majority of the civilised world was able to exist without a motorcar, and if asked if they would like one, most people would probably not have seen the utility of such a device.

The first requirement therefore must be to demonstrate what can be done and then to adapt the system as far as possible, over a familiarisation period, to suit the requirements which emerge from the user, after he has been exposed to the new technique. This therefore calls for an active approach, inasmuch as the providing service must go directly to the user. Here there is a difficulty, since if a new, highly technical, development, such as RECON, is available, it is rather difficult to persuade the user that this new tool is indeed as powerful as it is. It is conceivable that a touring demonstration unit, which links itself directly to the computer centre by dial telephone lines, could be established, on the lines of the touring NASA educational unit. In Europe of course, for an organisation covering 10 Member States the language problem becomes very important. The ESRO Space Documentation Service's answer to this situation was to set up its Customer Liaison Unit. This is a small team of five individuals, with occasionally the help of additional supporting staff. Their almost superhuman task is to cater for the disciplines covered by the NASA documentation system, in any one of the 10 Member State languages. I think you will realize that to do this job properly, almost certainly a much larger staff unit would be required. However, I think you will also agree that in this nucleus of five people, we have a rather unique approach to the problem. It can be regarded as a spearhead sales force with a portfolio of services, which often by the way have to be paid for, just as if it was a collection of detergent samples. Perhaps in a way the latter are easier to sell, particularly with the addition of an enzyme or two.

THE IMPORTANCE OF COST

A wide divergence of opinion seems to exist on the question of the correct price for a computer search (apart from the question of actual cost). A librarian user would invariably say the machine services are too expensive for my budget; a market research consultant will say the services are too cheap, and therefore they will not be appreciated at their true value. Somewhere in between lies the user, who if he has a real need would almost certainly pay whatever is a fair price.

We have found that it is easier to promote the use of mechanised services on a paid basis for retrospective searches of a large file, where the amount paid bears some relation to the true cost of the search. The user himself, can usually see the major advantages in being able to coordinate numbers of terms at a time, whilst searching a large collection. SDI services however, although extremely popular, prove more difficult on the rather more expensive tailor-made system. Thus emulating NASA, although in this case quite independently, we have found it necessary to introduce what we call a Standard Profile service, and NASA calls "SCAN". The Standard Profiles are obtainable at roughly 1/5 of the price of the tailor-made product. They are of course not tailor-made, but they are produced on a continuing basis and they clearly satisfy a customer requirement. The effect of the introduction of this new service this year, may be seen in the attached graph. It is an interesting fact that relevance levels need not be too high if the customer pays for the service, unless he pays per item. Most of them consider that a little "noise" is acceptable, even a little desirable, since it offers a limited facility, although admittedly a poor substitute, for browsing.

THE USER'S REQUIREMENTS

In general, when the user is paying for the services provided, the first requirement seems to be speed of response. In many cases even lists of accession numbers are preferred if they can be obtained quickly. Consequently, the argument is that with a basic list of references, the user can often immediately find some of the more important items for himself. Subsequently, if a comprehensive bibliographic listing follows the initial list of accession numbers, the user may make a second selection of the items of secondary importance, based on the fuller information provided.

The next requirement appears to be full coverage of available sources. It is fairly clear that many users would like to be able to go to one unique centre for documentation on any subject need. I believe that some documentation centres in Europe are working towards this ideal, but there are of course the usual difficulties with regard to mixed machine systems. The MARC system would appear to be one of the first steps in this direction, and we are watching this with considerable interest.

The third requirement is document availability. As most of us well know it is very little use whetting the user's appetite with long lists of unavailable publications. We ourselves are only able to provide copies of

certain NASA documents from the microfiche file deposited with ESRO by NASA. In this respect there is a definite lack in the completeness of the service we offer. However, you are familiar with all the difficulties connected with copyright and storage problems.

Finally cheapness of service is of importance for routine maintenance of users' information needs. Most people will pay large sums for urgently required material, if they are sure that the service is fast, and that they are getting as complete a coverage as possible. We have a peculiarly difficult situation, in as much as we provide both free and paid services at the same time to the same users. Currently we are thinking of changing our price arrangements so that only retrospective file searches are provided free of charge. Such searches would only be made available of course to institutes in member states actively engaged in the ESRO programme.

NEW SERVICES

A new service is emerging from a requirement first stated by our Establishment in Holland, the European Space Technology Centre (ESTEC). This is the ESRO Establishment where current satellites are designed and tested. The problem here is one of selecting from a large number of alternatives, as is for example necessary when choosing a material or component for a particular application. We are therefore now in the feasibility and design study stages for a new project which is the creation of a databank of information on electronic components, with the capability for extension to include other data fields such as, for example, materials for use in the space environment. Initially this will be a databank of quality and reliability test data, to be used as an aid in the selection of high-reliability electronic components for use in ESRO satellites. In addition however, we plan that the databank should contain a basic level of manufacturers' catalogue data, for all European electronic components.

Access to the data will be possible by conventional batch processing, which will provide for the generation of selective dissemination notices presenting detailed information on each component selected, and for search queries to be processed against the entire databank. In addition a remote on-line interactive interrogation system will be implemented utilising the same peripheral equipment as NASA-RECON. The input, display and printout of graphical data is being examined and will be incorporated at an early date. A third mode of operation is possible which is expected to be used extensively not only by SDS documentalists, but also by scientific and technical staff when copious output is expected. This is intermediate between batch-mode and on-line interactive mode and is a form of remote batch entry; profiles or search queries will be defined using the terminal and then stored for later processing in the batch stream on a priority system. A generalised system flow diagram is shown.

Clearly the databank now envisaged should be potentially very valuable not only to ESRO, but also to the entire European electronics industry. We intend that it should be fully exploited by making available a variety of services against subscription. These will include say, fortnightly current awareness notifications of newly introduced products in considerable detail. Later we would hope to produce quarterly and annual cumulative specific catalogues. Evidently also, a search against the entire data set would give a state-of-the-market survey.

CONCLUSIONS

In a paper entitled "Planning and Development of the European Space Documentation Service: an Example of International Collaboration" (Ref.3), the author stated, "It is at least in theory possible to imagine a group of countries deciding to create an information system related to several scientific and technological user-areas ..... there is a real prospect that with large third-generation computers a system of direct interrogation of the computer by users will become possible. The implications ..... require urgent study ....." This paper was written in 1967 and the conclusions are even more valid today. The NASA documentation system already covers a large number of disciplines, and based on it, we in ESRO, can provide a number of services, including direct access capabilities. The addition of databank features, also with direct access possibilities, increases the scope of services available. It is also conceivable that the NASA system will be applicable to other documentation files in the near future.

All the ingredients are there - is the user yet educated enough to act as the catalyst which produces the ultimate system?

REFERENCES

1. *Study of Mechanization in DOD Libraries and Information Centers.* AD 640 100. Booz, Allen Applied Research Incorporated, Bethesda, Maryland, BAARINC Report No. 914-1-1, September, 1966.

2. *Scientific and Technical Communication: A Pressing National Problem and Recommendations for Its Solution.* National Academy of Sciences, National Academy of Engineering, NAS Publication 1707, 1969, 336 pp.

3. Page, J.R.U. *Planning and Development of the European Space Documentation Service: An Example of International Collaboration.* Ciba Foundation Symposium on Communication in Science: Documentation and Automation. European Space Research Organisation, Paris, 1967, pp. 146-155.

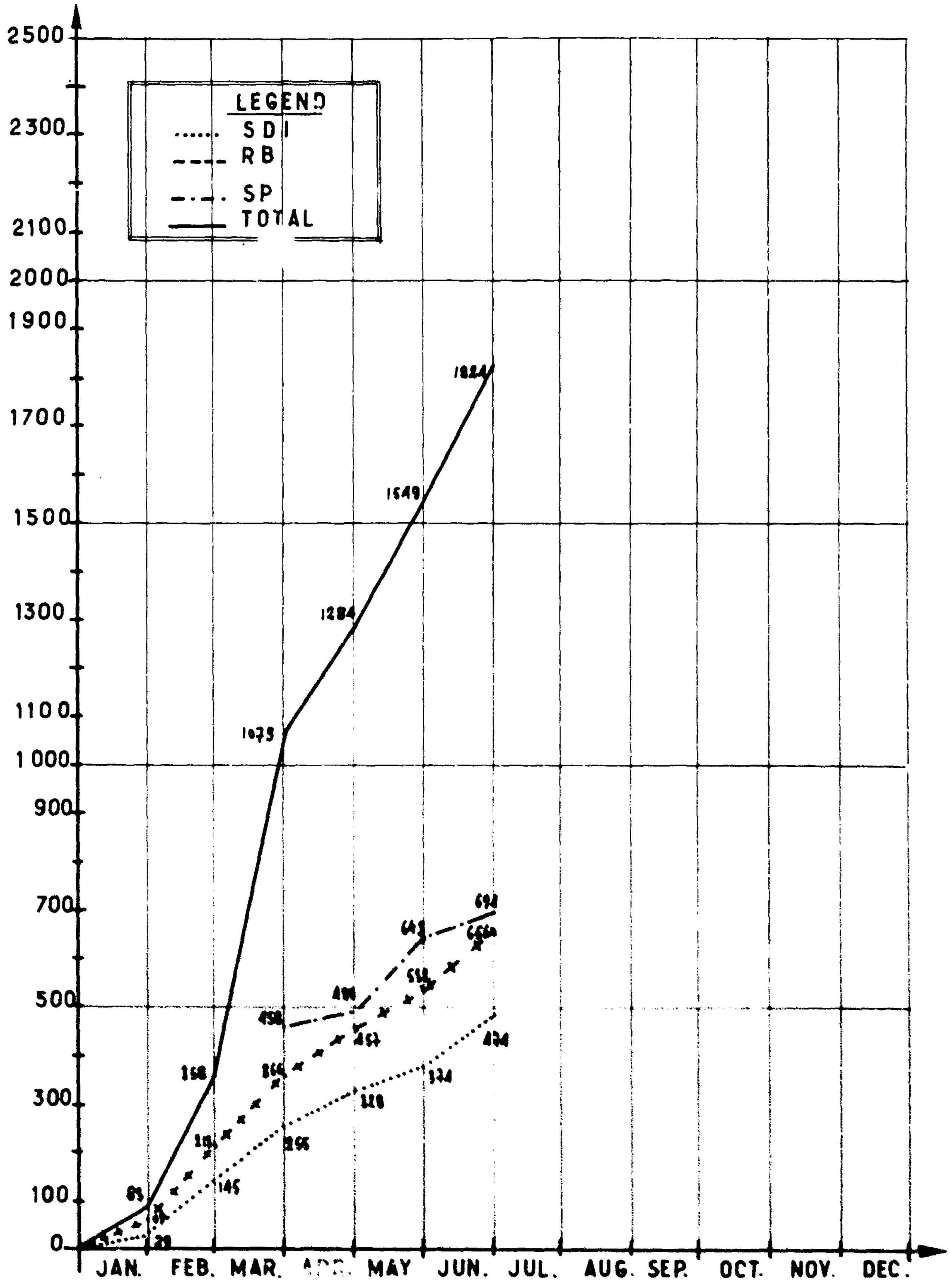


Fig. 1 Space Documentation Service bibliographic searches in 1969

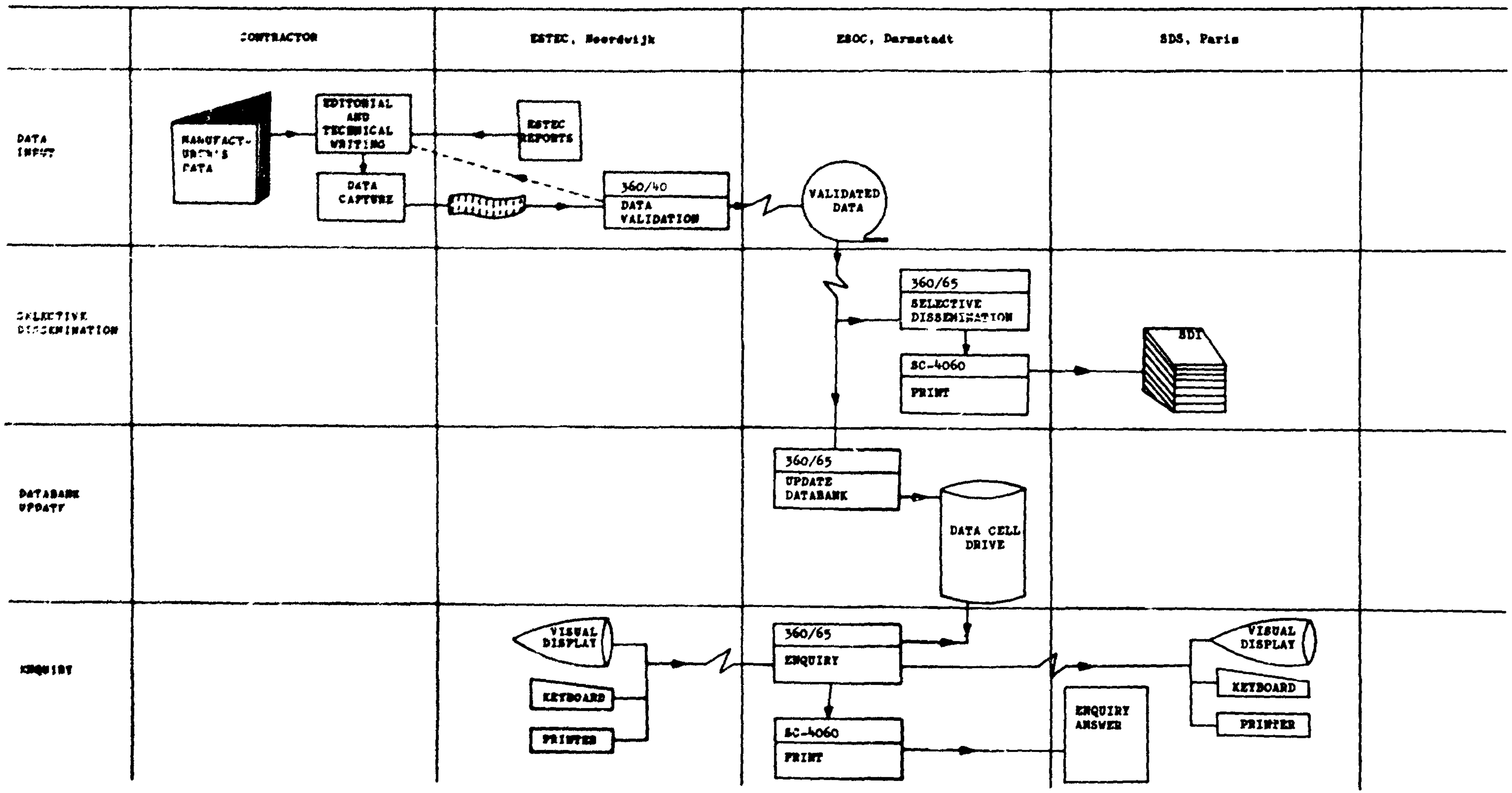


Fig. 2 Databank (electronic components and materials)

PAPER 4

PREPARING THE ORGANIZATION FOR AUTOMATION

by

Davis B. McCarn

Deputy Director, Lister Hill National Center for Biomedical Communications,
National Library of Medicine, Bethesda, Maryland, USA

SUMMARY

Careful planning and tactful action to ready the organization are critical to the success of automation. In spite of the obviousness of this platitude many egregious system failures have resulted from unrecognized organizational resistance. The assault of the system analysts and programmers is often either naive or machiavellian; naive when it assumes that computer technology is the wave of the future and no rational employee could object to the new system; machiavellian when technology is used as a pretext to wrest control of the organization's information network and, thus, control of the organization itself. To facilitate an effective automation, management must recognize these potentials for conflict; act to resolve those which develop; develop and maintain a climate hospitable to change; involve personnel throughout the organization in system design and development; recognize and act to reduce unrealistic uncertainties and fears; control the enthusiasms of the technologists; and insure that employees are trained to accept new ways and new roles.

"As the births of living creatures at first are ill-shapen, so are all Innovations, which are the births of time."

Bacon

SOMMAIRE

Une planification approfondie et des dispositions pleines de tact sont des atouts essentiels au succès de l'automatisation d'une organisation. En dépit de l'évidence de cette remarque, on enregistre des échecs nombreux et insignes résultant d'une résistance non reconnue de l'organisation. Les méthodes d'approche des analystes de systèmes et des programmeurs sont souvent naïves ou machiaveliques: naives lorsqu'elles posent en postulat que la technologie de l'ordinateur est la voie de l'avenir et que nul employé doué de raison ne peut émettre d'objection à l'égard du nouveau système; machiavéliques quant la technologie sert de prétexte pour obtenir par la force le contrôle du réseau d'information de l'organisation, et ainsi celui de l'organisation elle-même. Pour faciliter une automatisation et la rendre efficace, la direction doit reconnaître ces causes potentielles de conflit; prendre des dispositions pour résoudre ceux qui apparaissent; créer et maintenir un climat favorable au changement; faire en sorte que le personnel de toute l'organisation soit concerné par la mise au point et le développement du système; reconnaître certaines incertitudes ou craintes non réalistes, et s'efforcer de les réduire; modérer l'enthousiasme des technologues et s'assurer que les employés reçoivent la formation propre à leur faire accepter de nouvelles méthodes et de nouveaux rôles.

PREPARING THE ORGANIZATION FOR AUTOMATION

Davis B. McCarn

The preparation of this paper has been a humbling experience. A hurried review of fifteen years of computer systems development and administration revealed a singular gap with regard to organized thought on preparing the organization for computer systems. Upon further reflection it seemed clear that the problem was far from trivial, and, in fact, the majority of known system failures had probably resulted from man rather than machine problems. I had been certain that the textbooks and professional literature would provide the framework and strategy I had missed in my previous reading. I went through my two feet of old *Datamation* magazines; I found no articles at all! Somewhat shaken I turned to the *Computing Reviews* (Refs. 2-5) and searched 1960 to 1967 and found nothing. While I am not a good searcher and there may be articles I missed, the conclusion still seemed inescapable that the information technology professions in their professional literature were unconcerned about actually getting organizations to adopt their systems. But the problem is very real, and multiplying "user needs" studies bear mute testimony to one aspect of it.

As an example of a typical problem, I had occasion not long ago in a previous job to observe the floundering of a new system development effort, a large-scale, integrated, on-line, information handling system for an organization which already had a successful batch processing system. The batch processing system, itself, was in trouble because the organization had had it built completely by contract and had allowed the knowledge of this system to be concentrated into the hands of one or two individuals. These individuals had finally left the organization, and it was without any knowledgeable programmers or documentation on its operational system, one which had to be run on a regular 24-hour-a-day basis to provide essential services to users. Thus, the organization was trying to simultaneously build a new system and survive the crises of the old.

The *leading actors* in the development of the new system were the chief of the computer center and the chief of the division which the new system was supposed to serve. Their relationship was star-crossed from the beginning, because the computer center had once been controlled by the division chief for whom the new system was being built. It had, however, recently been moved into a higher organizational position so as to be able "to serve other elements of the organization".

The system to be built was not well-defined. First, its goal was clearly to automate, for the sake of automation, the input process to the old system. There had never been an analysis that indicated that such automation was desirable, nor that it would produce a more efficient or a more economic operation. The faith certainly existed, however, that it would somehow make it possible to meet the rising demands for services for which additional personnel could not be obtained. It also had the major advantage of being non-threatening to the top managers in the operating elements because the new on-line system would support their internal effort rather than change the nature of their relationship to each other or to the users of the services of the organization. A major software house had the audacity to accept a contract to build the new on-line computer system and was soon trapped in the cross fire between the principals and their staffs.

Since the system was not well-defined, the first step was to define clearly the functions and operations that it was to perform. The using division, however, felt very threatened by the insistence that it define its requirements with precision. The specter that haunted it was of having some set of requirements finally accepted; having the computer people go off and build the system to meet those requirements; and then having them come back with a system that met every single one of the precisely-stated requirements but failed to help accomplish the floundering functions of the operating manager.

This threat, of course, was not illusory. Such systems have been built, and such results have occurred. In this particular case, a cycle of definition and re-definition, using more and more abusive language and involving more and more abrasive encounters, developed between the two staffs with the contractor as the shuttlecock.

As the situation worsened, the operating division manager finally forbade all his people to talk to the contractor or the representatives of the computer division, and the head of the computer division refused to meet with the head of the operating division. This impasse had to be resolved by higher authority. I am sure you will be glad to know that the story had a happy middle; the internal factions were momentarily resolved; the organization joined forces against the contractor; and a stiff penalty was imposed on him for failing to deliver the initial phase of the system in time. It still remains to complete the whole system; only the first round is over.

The National Library of Medicine has had similar problems. It began seriously considering automation in 1949 and moved into a punch-card/Listomatic system in 1959. This original system was also born in a time of scarcity. The Library had had no increases in authorized personnel for ten years, in spite of having graduated out of the Army to become a national resource. A singularly candid account (Ref. 14) of this development reads in part:

"January 1960. A catastrophe was discovered on January 5. For an unknown length of time the camera had been operating with no exposure light flashing. The entire subject and author sections had to be rerun. Page stripping was performed on a crash basis by four operators.... The Indexing Assistant backlog rose to 6,300; the input typing backlog rose to 4,200; the proofreading backlog rose to 3,455; the keypunching backlog dropped almost 3,000 to 584; the output imprinting backlog rose to 3,208. Input typewriters suffered from minor problems such as blown fuses. Keypunching improved. Output Justowriters were proving to be prone to breakdown; there were eight days of down-time here. Adequate film processing in Washington was still unavailable, and film was transported via Trailways Bus delivery service to be processed in New York."

It seems evident that some of the woes of this system may have resulted from resistances in the organization. This system paved the way, however, for the present computer system, the Medical Literature Analysis and Retrieval System (MEDLARS) (Ref. 1). Thus, the NLM advanced by increments to its present system; an option no longer open.

Were these experiences isolated ones? A review of literature found through other sources appears to indicate that they were not isolated phenomena. McKinsey and Company ran a survey (Ref. 11) in 1963 of some 27 companies and identified three factors as being crucial to the few successful computer systems. These were: executive leadership, management control, and the involvement of operating management.

The same company repeated the survey (Ref. 12) in 1968 and concluded gloomily, "from a profit standpoint, our findings indicate, computer efforts in all but a few exceptional companies are in real, if often unacknowledged, trouble". The major problem identified by the survey was the gap between the computer staff and the operating management. To bridge that gap, McKinsey and Company proposed: improved teamwork in the corporation between the top management, operating managers, and the computer professional; more extensive reliance on evolutionary development of computer systems rather than turn-key operations; careful consideration and use of communications with computers and on-line, remote-access systems; and, most important of all, improved leadership from the top executives in the company.

Another recent survey (Ref. 18) done independently by Putnam reviewed the success of computer systems in some 20 films and found the major sources of resistance to integrated information systems in the large corporations to be: lazy senior management; insecure operating managers who are threatened by automation; organizations which are themselves already sick and are more concerned with being able to affix blame for efforts than credit for effective action; executives who are "egocentric" and sufficiently powerful to override the valid objectives of the organization; and politics within the company. The most difficult problem with regard to installing unwelcome systems was politics, because the manager who is acting from political interest can be completely rational and most effective in his action to defeat the success of a new system. Such activity is at its worst in organizations which have internal empires ruled by their own feudal lords. The suggestions of Putnam to solve these difficulties and dangers were: involvement of the senior company executive; obtaining outside, objective evaluation judgment (from a management consultant, naturally); and a constant alertness to the human problems involved in the systems implementation.

It would certainly seem from this brief review of available surveys that the human problem in the organization deserves careful consideration in planning the installation of a new system.

There has been substantial work on innovation in industrial organizations. The business administration world has been studying it in their schools for many years. There are good texts on organizational behavior (Ref. 8). The Institute for Research on Human Behavior has charted some of the behavioral science aspects of the problems (Refs. 10 and 17), and Mumford (Ref. 16) has summarized much of this work in relation to EDP in a very perceptive article. She pointed out that the stability of the organization, the user perception of the change, the strategy for change, and the role perception of the computer group are the important parameters in determining the success or failure of a system.

My own experience leads me to believe that the installation of new systems begins in a paradox, also identified by Mrs Mumford: a stable organization satisfied with its own performance and providing personal satisfaction to its employees cannot be automated, but automation as a cure in an unstable organization is fated to be born in confusion and conflict. Many attempts to introduce ADP into stable, well-functioning organizations have met with open resistance or apathy. As Michael (Ref. 15) observed, "The risks of individual, interpersonal, and organizational failure involved in deliberately accepting the uncertainty that must accompany organizational and individual changes are usually too great for men to take willingly when things are going well and when men and institutions have been successful.... The evidence indicates that basic organizational changes can occur through disasters or by deliberate organizational change programs. The latter must be directly and continuously supported from the top of the organization over many years of unremitting and highly organized effort. Such deliberate efforts have been rare....." Thus, one could conclude that the efforts which have been undertaken must have largely been spurred by the threat of disaster.

Now the disaster which threatens information centers is overload of input and insatiable demand for services. As such overloads develop, compensating actions are taken by th[illegible]ization. These are described in a study of a university library by Meier (Ref. 13). As loads increase [illegible] efficiency eventually drops, much as the throughput of a highway drops with heavy traffic. Stress and [illegible] increase with backlogs and complaints until collapse threatens. (Systems *do* collapse, they don't just fade away. The Chicago mail system collapsed several years ago.)

As the specter of collapse begins to materialize, the organization may turn to thoughts of ADP. The hope being, of course, that ADP will stave off disaster. In reality the installation of ADP systems requires investment of scarce resources in design and in parallel operation which further degrade the performance of the information center. But ADP development offers *hope* of relief. If that hope is kept within reasonable levels and if the system progresses so as to maintain hope, a success may result. If expectations become excessive or the system development stumbles, the organization will lose faith and the system will fail. As Hoffer observed in *The Ordeal of Change* (Ref. 7) "When a population undergoing drastic change is without abundant opportunities for individual action and self-advancement, it develops a hunger for faith, pride, and unity. It becomes receptive to all manner of proselytizing, and is eager to throw itself into collective undertakings which aim at 'showing the world'. In other words, drastic change, under certain conditions, creates a proclivity for fanatical attitudes, united action, and spectacular manifestations of flouting and defiance; it creates an atmosphere of revolution." If these energies can be channeled toward the development of a new system great progress is possible; if they oppose it failure is sure.

Thus, often the first problem of the system developer is the harnessing of irrational forces to provide the extraordinary effort required. At the same time, however, he must move to increase the rationality of the organization. This rationality can be improved in a variety of ways. The first is to reduce the uncertainties and fears associated with the new system and move to avoid excessive expectation. These objectives can be achieved through preparation of an implementation plan and the presentation of that plan to affected members of the organization. Such a plan should justify the new system and specify how it will resolve the crises or inadequacies of the present system. It should provide an overview of the new system and of how it will help the function it is to support. The plan should specify the probable impact on organizational elements, managers, and individuals, and provide people throughout the organization with information on job changes, training, and replacement programs. It should provide a schedule for the effort and specify the costs. The schedule should be clear on the period of parallel operation. Finally, it should specify who or what organizational elements are expected to participate and what the responsibility of each is. The press of time and the rate of change of some aspects of the above may make it difficult to produce a formal plan but such information should exist and be available to the concerned staff.

Not to provide such information raises the concern and resistance of personnel who may believe they are confronted with a machiavellian manipulation. Cooperation and rationality must be fostered through clear specification of goals and the road to these.

Cooperation however will not be enough. Developers of systems must continuously strive to support their development through legitimate power. As every survey found, top management must support the system. Maintaining this support when the organization is faced with day-to-day operational problems is not easy. but without such support the human investment for the "Great Leap Forward" will not be found in time, momentum will flag, and the movement will fail. It is particularly important that the reward and punishment system reflect this top management support: Those who contribute to the new system should be rewarded and promoted, rather than dropped from the mainstream of advancement as is sometimes the case. Similarly, effective deterrents to resistance need to be developed.

To provide a vital incentive to the development, it should be a little daring. Doing what many others have done before because it looks safe may not be safe. A trite system may look easy to build; but it will not mobilize effort the way an imaginative one can. Pride in technological sophistication can be an important stimulus to all levels toward effective systems development, and it has become a goal in itself of some corporations (Ref. 6).

Harnessing irrationality and creating the climate for change must be paralleled by efforts to involve all levels of the organization in the development. Managers must be involved in establishing goals, directing and assisting in functional analyses, and explaining the system to their staff, users, and outside groups. This new system should not become the property of the ADP staff, but all management must assume a stake in the success of the effort.

Of almost equal importance to the managers are the opinion leaders. These must be converted to "change agents". They should be involved in the preparation of requests for proposals, evaluation of proposals, selection of hardware and software contractors, functional analyses, etc. It is particularly important that these people see equipment and comparable systems. Those who will have major roles in the new system may be detailed to work with the developers.

Other staff to be affected need education in the concept of the new system, information on the plans for its development, and encouragement to help carry the extra load the development entails. They also need to be realistically appraised of the "unforgiving" nature of computers with regard to sloppy inputs, new quality control expectations, and new procedures. Some detailed planning should be done at the work group level. As C.P. Snow (Ref. 19) noted with regard to Tizard's installation of Radar, "To get anything done by any highly articulated organization, you have got to carry people at all sorts of levels. It is their decisions, their acquiescence or enthusiasm (above all, the absence of their passive resistance), which are going to decide whether a strategy goes through in time."

The new system will probably involve high levels of conflict, either with large segments of the organization, with informal work groups, or with individuals. An attempt to "paper over" these conflicts, an attempt to solve or resolve them by dictatorial edict or other efforts that do not contain the element of negotiation are likely to result in major problems in the systems development. The signs of conflict are usually relatively

clear, if one is willing to see them. Persistent failures to agree, delays in acting on important paperwork in the systems development, refusal to provide inputs, the maintenance of parallel systems generally with better information, and refusal to use a new system are all obvious resistances. Less obvious perhaps are those kinds of resistance that attempt to kill a system with kindness: the insistence that it move promptly from research and development into operation is only too common, as is the insistence on rapid changes even before operation of the system. Such conflicts must be faced and resolved if the system is to succeed.

This description had focused on the problems of the developer; it is only fair to note that the future user may also have *realistic* objections to the new system. Technological enthusiasms, power hunger, know-it-all attitudes, and small errors on the part of the developers also must be avoided. Another paper could easily be written for users on how to avoid unrealistic and rigid systems, but that would be for a different audience and about some other developers.

In any event the way of the reformer is not easy. As Machiavelli (Ref. 9) observed half a century ago, "It must be considered that there is nothing more difficult to carry out, nor more doubtful of success, nor more dangerous to handle, than to initiate a new order of things. For the reformer has enemies in all those who profit by the old order, and only lukewarm defenders in all those who would profit by the new order, this lukewarmness arising partly from fear of their adversaries, who have the laws in their favour; and partly from the incredulity of mankind, who do not truly believe in anything new until they have had actual experience of it. Thus, it arises that on every opportunity for attacking the reformer, his opponents do so with the zeal of partisans, the others only defend him half-heartedly, so that between them he runs great danger. It is necessary, however, in order to investigate thoroughly this question, to examine whether these innovators are independent, or whether they depend upon others, that is to say, whether in order to carry out their designs they have to entreat or are able to compel. In the first case they invariably succeed ill, and accomplish nothing; but when they can depend on their own strength and are able to use force, they rarely fail. Thus, it comes about that all armed prophets have conquered and unarmed ones failed; for besides what has been already said, the character of people varies, and it is easy to persuade them of a thing, but difficult to keep them in that persuasion."

REFERENCES

1. Austin, Charles J. — *MEDLARS 1963-1967*. National Library of Medicine, Bethesda, Maryland.

2. — *Comprehensive Bibliography of Computing Literature; Permuted and Subject Index to Computing Reviews, 1966*. Association for Computing Machinery, New York, 1966.

3. — *Permuted and Subject Index to Computing Reviews 1964-1965*. Association for Computing Machinery, New York, 1965.

4. — *Comprehensive Bibliography of Computing Literature; Permuted and Subject Index to Computing Reviews, 1967*. Association for Computing Machinery, New York, 1967.

5. — *Permuted (KWIC) Index to Computing Reviews, 1960-1963*. Association for Computing Machinery, New York, 1963.

6. Galbraith, J.K. — *The New Industrial State*. Houghton Mifflin Co., Boston, 1967.

7. Hoffer, E. — *The Ordeal of Change*. Harper and Row, New York, 1952, pp. 4-5.

8. Lawrence, Paul R.; Seiler, John A. — *Organizational Behavior and Administration: Cases, Concepts, and Research Findings*. Rev. ed. Richard D. Irwin, Inc., Homewood, Illinois, and The Dorsey Press, 1965.

9. Machiavelli, Niccolo — *The Prince and The Discourses*. Random House, New York, 1950.

10. Mann, F.C.; Neff, F.W. — *Managing Major Change in Organizations*. The Foundation for Research on Human Behavior, Ann Arbor, Michigan, 1961.

11. — *Getting the Most Out of Your Computer: A Survey of Company Approaches and Results*. McKinsey & Company, Inc., New York, 1963.

12. — *Unlocking the Computer's Profit Potential*. Computers and Automation, McKinsey & Company, Inc., April 1969.

13. Meier, Richard L. — *Social Change in Communications-Oriented Institutions*. University of Michigan, Ann Arbor, Michigan, March 1961.

14. Miller, E.A. — *The National Library of Medicine Index Mechanization Project*. Bulletin of the Medical Library Association, Vol. 49, Part 2, January 1961.

15. Michael, D.N. Science, Vol. 165, 11 July, 1969, pp. 165-6.

16. Mumford, E. *Implementing EDP Systems: A Sociological Perspective.* The Computer Bulletin, January 1969.

17. Peter, H.W. *Guidelines in the Process of Change.* In: International Cocperation and Problems of Transfer and Adaptation, Vol. X. United States Papers prepared for the United Nations Conference on the Application of Science and Technology for the Benefit of the Less Developed Areas, US Government Printing Office, Washington, DC, 1963.

18. Putnam, Arnold O. *The Human Side of Management Systems.* Business Automation, Vol. 15, No. 11, Nov. 1968.

19. Snow, C.P. *Science and Government.* Oxford University Press, London, 1961.

PAPER 5

FINDING THE OPTIMUM MIX OF STAFF AND CONSULTANTS

by

G. X. Amey

Defence Research Board,
Department of National Defence, Ottawa, Canada

SUMMARY

The problem of staffing an information service during the transition from the manual to an automated system is analyzed. The stages of development are first presented in detail so as to indicate the re-education of all levels of staff as they begin to realize the pay-offs possible from such a system and the demands it will make of its exploiters if the possibilities are to be realized.

Of the development strategies possible, DSIS chose to emphasize that in which the work is carried out co-operatively between the internal staff and consultants external to the organization. This helps to minimize difficulties with internal staff, by ensuring the involvement at all times so that they feel the system being developed is their own. To optimize the man-machine system it is necessary to restructure the organization to make best use of the special qualities of each of the principal elements. Methods are practical in small manual systems in which there are no true specialists and everyone does a little of everything. This approach must be rationalized during automation to ensure that there are no weak links in the system.

After the preliminary stage in which the general concepts were tested inhouse, there have been about equal numbers of computer-related staff in DSIS and supplied by contractors, though the balance of expertise is with the latter.

SOMMAIRE

L'auteur analyse le problème qui consiste à pourvoir en personnel un service d'information pendant la phase de transition entre un système manuel et un système automatisé. Il présente d'abord en détail les étapes d'un tel développement, de façon à donner une idée du recyclage du personnel à tous les niveaux, lorsque celui-ci commence à prendre conscience des avantages possibles offerts par un système de ce genre, et des exigences imposées à ceux qui l'exploitent si ces possibilités se concrétisent.

Des différentes stratégies possible, le DSIS a choisi, pour la mettre en relief, celle où le travail est effectué en coopération par le personnel intérieur et des consultants extérieurs à l'organisation. Cette solution aide à réduire au minimum les difficultés survenant avec le personnel intérieur, en assurant a tout moment leur implication et, par conséquent, en créant chez eux le sentiment que le système mis au point leur appartient. Si l'on veut perfectionner le système homme-machine, il faut restructurer l'organisation pour utiliser au mieux les qualités spécifiques de chacun des éléments principaux. Les méthodes sont pratiques dans ce petits systèmes manuels pour lesquels il n'ya pas de véritables spécialistes, et où chacun fait un peu de tout. Cette méthode d'approche peut être rationalisée en cours d'automatisation pour éviter que le système ne comporte des maillons faibles.

Après le stade préliminaire, au cours duquel les concepts généraux ont été mis à l'essai au sein de la société, on s'est trouvé en face d'un nombre à peu près égal d'employés appartenant à la DSIS et de personnel sous contrat, bien que le bagage technique de ces derniers ait été supérieur.

FINDING THE OPTIMUM MIX OF STAFF AND CONSULTANTS

G.X. Amey

1. THE PROBLEM

It is desired to transform the staffing of a manual documentation system into that required by an information system based on the use of computers. This calls for essential changes in the staff organization as well as in the total number of staff during each phase of the transition. In order to trace these changes within the system, the growth and development of the system itself must be considered. This is covered fully in my paper "Evolutionary Transition to Automation of Information Management" (Ref. 1).

Evolution of the System

The transformation of the system is usually divided into a number of stages in each of which, machine-interaction increases and staff-roles are modified to make more effective use of the computer as understanding of its merits and shortcomings are appreciated. Each stage will tend to continue in parallel with the new stage for some time during a transition period as shown in Figure 1.

In the first stage, the computer involvement is likely to be as a mere addendum to the system. Since the formation of a data base is a prime requisite for operation of an automatic retrieval system, this task may be undertaken long before there is any possibility of extracting benefits from the system. This stage also brings the painful realization of how important absolute accuracy is in dealing with a machine. Our own experience provides an excellent example of improper use of computers in which just one element of the manual system is replaced by a mechanical equivalent. We initially developed a method for automatic selective dissemination of catalogue cards printed on IBM tab cards. This proved to be more cumbersome than the manual sorting operation currently used so that the method was not worth implementing.

In the second stage of development, we have made good use of the capabilities of computers for sorting and rearranging data in formatting and indexing our announcement services. This allowed some small savings of clerical effort but the principal payoff is a reduction of effort in processing classified documents plus the addition of indexes to our announcements. In computer terms this is a batched operation which will not be obsoleted by further development of the system.

We shall finally reach a stage where the machine is fully integrated into the system and acts as a symbiont of the staff. In Figure 2 DSIS is a black box, BB2, which receives inputs from suppliers of information and provides outputs to final users. Neither the source or the user need concern itself with the internal operation of the black box. Within DSIS there is a second black box, BB1, the online automatic system which accepts a variety of inputs and provides complex outputs according to set rules. The evolution of this black box determines the changing requirements of the staff interacting with it.

2. STRATEGIES OF STAFF EVOLUTION

Turnkey Systems - Minimal Involvement of Internal Staff

The most attractive approach in terms of economics and avoidance of administrative headaches, is to buy a turnkey system. This implies that somebody has already developed a system which appears to fit your requirements. In the most extreme case, machine, system and staff are supplied as a package, as in the case of NASA, where only two members of NASA staff are needed to oversee daily operations of the information services. A much larger staff is of course required to define and continuously evaluate such a system. The NASA system was tailored to their requirements and is an excellent example of an operational large scale system. More usually the problem is to convert an existing manual system and there is no turnkey system available, which is tailored to the specific needs of the user. Painful and expensive modifications are required to make the two systems fit.

Systems of this kind which are almost entirely developed by computer-oriented people, betray their origin by provision of services and requiring inputs which are essentially computer-oriented rather than directed towards needs of the ultimate users of the system.

Characteristics of such systems include the following: rigid definition and formating of inputting; report generators which produce only tabulations; use of printout in upper case only; provision of cryptic messages to the user, and over-reliance on the system to make what are essentially intellectual decisions. The system is essentially a black box with an elegant internal structure but unfitted in both input and output to human users.

The last element of computer performance is extracted from the system by bit-diddling and tailoring the system to the particular configuration used. Expansion and development of the system is very difficult and transfer to a different make of computer is impossible.

Development Shared by Internal Staff and Consultants

The more usual route is to involve internal staff heavily in the design of the system, using external consultants to provide specialized computer know-how. We are then confronted with the problem of deciding the degree of involvement of internal staff in this procedure in order to facilitate communication between them and the consultants. Your staff must become acquainted with computer technology and the consultants must become experts in your system. Provision of inhouse expertise in the computer fields means either your staff must be retrained or new staff with the required skills must be hired. Addition of staff goes contrary to the rationale of using computers - which are generally claimed to reduce staff. Retraining pre-supposes that current staff are capable of absorbing computer skills at a speed sufficient to enable continuing usefulness in the organization. Normal production must continue despite diversion of effort to development activity.

Contribution of the System User to Design

The third strategy in the development of a system is that of involving the user in the system. This appears necessary if it is to supply his needs; otherwise the system will probably be optimized from the viewpoint of the information staff. To involve the users in the SOCRATES system, annual conferences of librarians and information scientists have been held since 1966; evaluation questionnaires have been employed and the information scientists visit their main aggregations of users across Canada each year. More direct involvement in detailed design is difficult. Experiments in automated retrieval have been conducted at several establishments (Ref. 2), and one of our more sophisticated users, after developing an SDI system for CAS tapes has become head of the computer division of DSIS.

3. INVOLVEMENT OF INTERNAL STAFF IN SYSTEM DESIGN

We chose the second method described above. The introduction of an automated system will have direct effects on the policy and organization inherent in the existing system.

Impact on Policy

Policy in relation to reproduction had to be reconsidered at DSIS. A system which has a response of several weeks to a request for information, is tolerable in a purely manual system. In an online system, where response to a query to the computer is a matter of seconds, the mechanism for supplying documents must also be fast. At DSIS it is planned to supplant hard copy in favor of microfiche since the latter can be reproduced much cheaper and faster. Thus one day turn-around to a request should be possible if microfiche is supplied by DSIS. Storage of microfiche at DRB establishments will give still faster response to local users.

Organizational Changes

The system based on specialization of subject rather than function is at a disadvantage when associated with machines. If the information scientist normally indexes and abstracts documents, prepares and updates profiles, conducts searches, etc., within a narrow subject area, then his absence for several weeks will introduce delays of this order into a system that is designed to respond in seconds.

In addition, specialization in subjects rather than functions, inevitably causes subjective indexing of material, which in turn makes much of the material difficult to retrieve by anyone other than the original indexer. This is intolerable in an online system where the user can browse directly in the surrogate files.

New categories of staff will come into existence. If no explicit planning group was in existence previously one should be formed in order to define long range goals and provide an interface between outside consultants and internal staff. This group will specify a system and control and evaluate the work carried out by external consultants. New divisions, concerned with operation of a computer and its peripherals, must be defined and staffed.

4. STAFFING EXPERIENCE AT DSIS

The total staff of 70 (including publication) in 1966 has not changed in the interim, although several new positions have been added. Figure 3 shows the five year projection of staff requirements used in selling the plan to management. The rising straight line indicates extrapolation of expected work load in terms of input of information. The upper line indicates total staff requirements during the transition and the lower line shows growth of computer operating staff internal to DSIS. We have now added five of the seven computer staff planned, and all were recruited inhouse, either by retraining in DSIS or by transfer from some other part of DRB. Detailed changes of non-computer staff are shown below.

Inputting

(a) Descriptive cataloguing now requires rigid definition of the nearly fifty categories of bibliographic data defined. A member of this staff is now employed in proofreading data to be entered on magnetic tape.

(b) The tasks have been rationalized so that particular girls carry out specific duties, rather than each person doing a little of each type of task. Backup provisions are made to cope with the problem of absence. The system is planned so that inputting has preserved the same form from the time the data was typed on work cards, through the stage (now being phased out) of producing paper tape on flexowriter, to the current stage of online entry of data on magnetic tape.

(c) Analytic cataloguing has been modified only to the extent of replacing the subject headings of the original DSIS system by TEST descriptors and COSATI categories. Deeper indexing and unlimited abstract length are now practical. There has been a 25% saving in documents processed for the information scientists by the use of the KWOC index to announce unclassified material which is well-indexed elsewhere.

Outputting

The main outputs of the system are printed announcements, and documents distributed either on request, or on a selective basis as received. Automatic formatting and paging of the announcement service saves some effort in the print shop. Internal production of microfiche may eventually call for increase of staff in the micro-photography section.

5. THE CONTRIBUTION OF CONTRACT STAFF

Dammers (Ref.3) has found that it is better to do all development internally, retraining one's own staff to carry out new tasks. This has not been possible at DSIS and we have relied on consultants to carry most of the design load. We have tried both of the approaches to contracting which were discussed earlier, functional specification of a system which is then implemented by us (Ref.4); and the use of consultants as an extension of our own staff (Refs.5,6). The latter method has been more satisfactory, since the consultants then interact directly at all levels of DSIS staff to ensure that impractical ideas are not incorporated.

Contract Control Documentation

Our method of job specification is designed to provide thorough documentation of the work being performed at all stages of system analysis, coding, debugging and implementation. A conceptual analysis is first made by DSIS planning staff of the problems to be tackled in any given fiscal year, in accordance with the five-year plan. The Project Officer receives a job description for each of the tasks involved, with estimates of the effort required and expected completion dates. Priorities are then set so as to ensure smooth operation of the current phase of the system while introducing improvements on the basis of experience. Changes in agreed schedules have to be approved by DSIS.

Fitting development work into the fiscal year has been our greatest headache. Changing government policy with respect to use of contractors as extensions of one's own staff, has contributed to our difficulties. However, we feel that we have made the best possible use of our own resources by this contribution of contractors and internal staff.

REFERENCES

1. Amey, G.X. — *Evolutionary Transition to Automation of Information Management*. DSIS-R3, 1969.

   and

   Amey, G.X. — *SOCRATES - An Information System for the Canadian Defence Community*. DR-196, 1969, p 18-25.

2. Currie, D.J. — *Information Retrieval for the Scientist - a DRES Experiment*. Defence Research Board Report DR-196, 1969, p.5-12

   and

   McIvor, R.A. — *Experiences in Information Retrieval Using Chemical Abstracts*. Defence Research Board Report DR-196 1969, p.13-17

3. Dammers, H.F. — *Phased Transition from Non-Mechanized Information Storage and Retrieval to On-Line Computer Operation*. FID-IFIP Conference, Rome, 1967

4. Brown, W.P.I. — *An Experimental System of Automated Services for the Defence Scientific Information Services*. Contract 2PB6-199, DSIS 67-10769

5. Moxham, J.G. — *File Design for Project Socrates*. Contractor Report in DR-196, 1969, p 26-35

6. Grant, R.W. — *Communication Facilities for an On-Line Information Retrieval System*. Contractor Report in DR-196, 1969, p.36-46

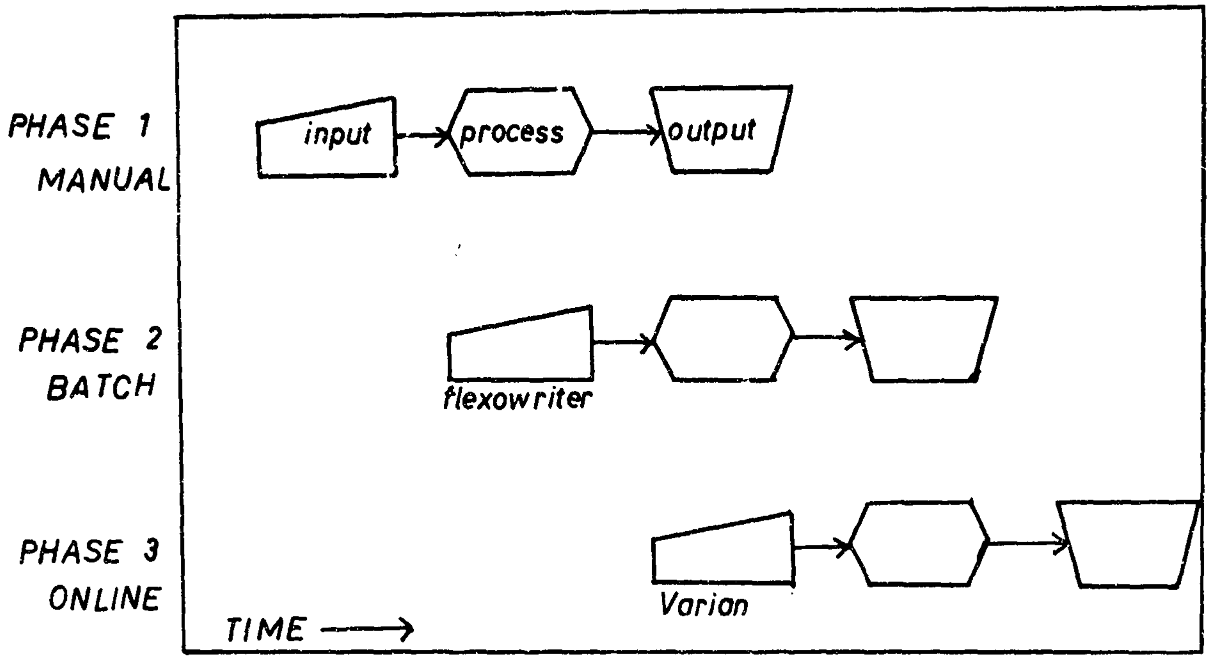


Figure 1

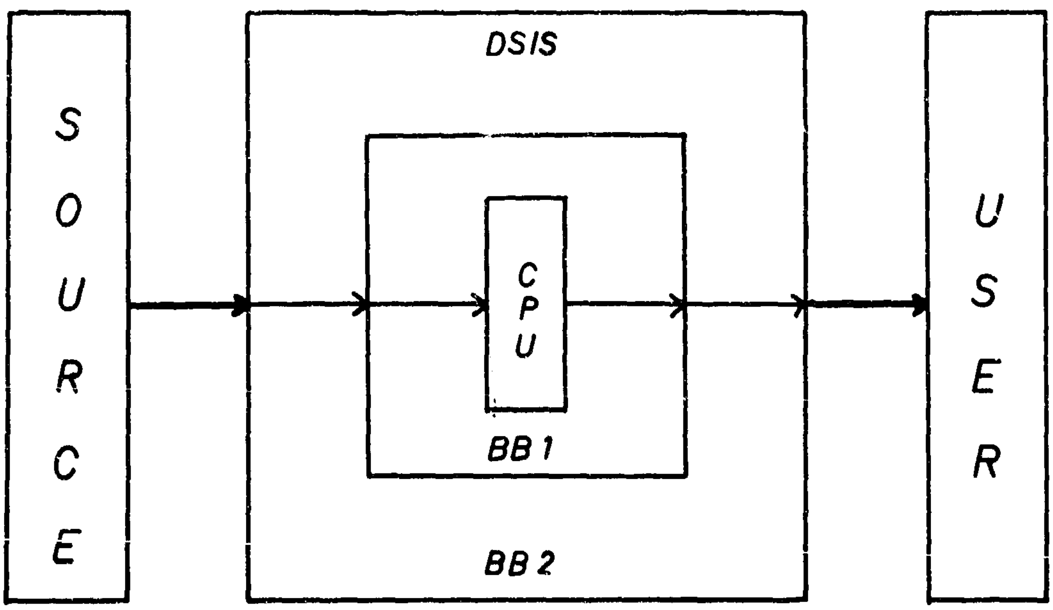


Figure 2

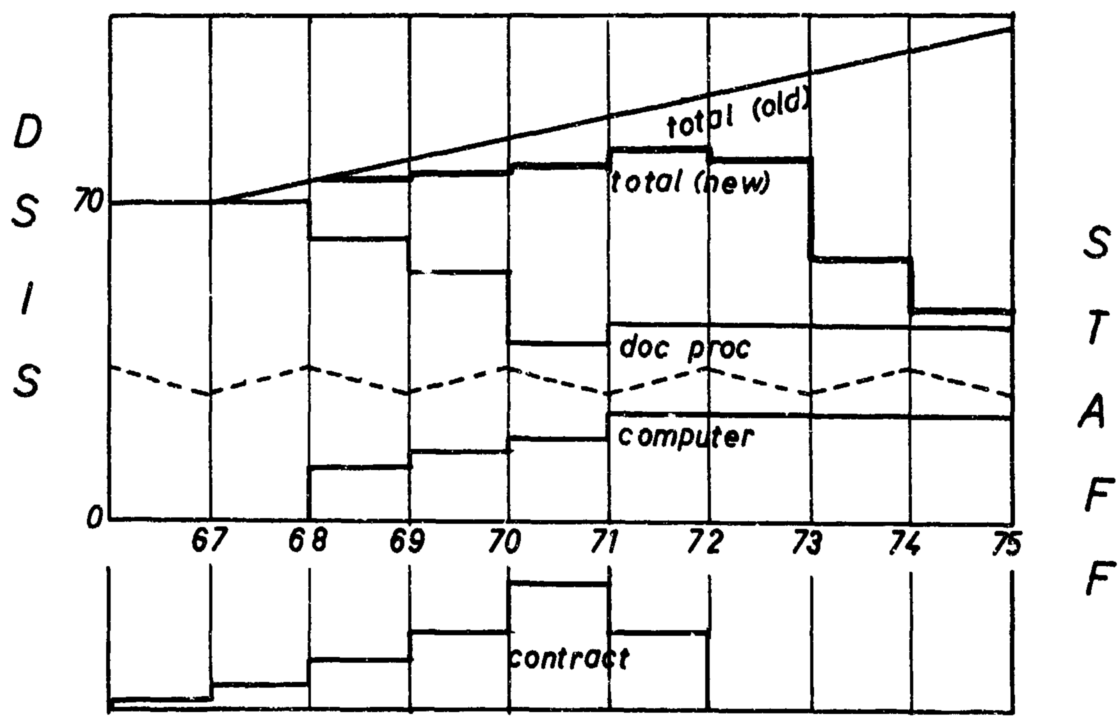

D
S
I
S
70
0
total (old)
total (new)
doc proc
computer
67
68
69
70
71
72
73
74
75
contract
S
T
A
F
F

Figure 3

PAPER 6

AUTOMATION OF THE INFORMATION SERVICE IN AN INDUSTRIAL RESEARCH ESTABLISHMENT: HARDWARE AND ITS EFFECT ON IMPLEMENTATION AND PERFORMANCE

by

H.F.Dammers

Shell Research Limited, Woodstock Agricultural Research Centre,
Sittingbourne, Kent, UK

SUMMARY

A comment is presented on hardware choice and its problems in the light of the experience gained in the development of computerized information facilities at the Shell Research Laboratories, Sittingbourne, UK. The environment in which the system operates is outlined as well as the objectives of the project and resources available. Computer configuration for data storage and retrieval operation at the site and links to other computers in the Company are described and the cost/effectiveness ratio of the different operations is established. Finally future needs and developments are discussed.

SOMMAIRE

L'auteur étudie le choix de l'équipement et les problèmes qu'il soulève, à la lumière de l'expérience acquise au cours de la mise sur pied de services d'information équipés d'un ordinateur, aux laboratoires de recherche de la SHELL, à Sittingbourne, Grande Bretagne. Il donne un aperçu de l'environnement du système ainsi que des objectifs du projet et des ressources disponibles. Il décrit ensuite la configuration d'ordinateur propre aux opérations de mise en mémoire et de localisation d'informations sur place, ainsi que les liaisons avec les autres ordinateurs de la Compagnie. Il détermine la rentabilité des différentes opérations et conclut en analysant les besoins et les développements de l'avenir.

AUTOMATION OF THE INFORMATION SERVICE IN AN INDUSTRIAL RESEARCH ESTABLISHMENT: HARDWARE AND ITS EFFECT ON IMPLEMENTATION AND PERFORMANCE

H. F. Dammers

1. INTRODUCTION

The attitude of computer users and operating staff with regard to hardware matters tends to be a rather ambivalent one. On the one hand, one can validly take the line that hardware matters are of relatively little consequence, as with our present facilities the main bottlenecks tend to be in the software area. On the other hand, constraints imposed by our available hardware capabilities or incapabilities are often all too evident and constitute not infrequently a source of severe frustration.

Hardware limitations were probably the major barriers to overcome in the application of computers to practical information storage and retrieval systems, and it is perhaps fair to state that only with the arrival of second generation computers, such as the IBM 1401 early in the sixties which combined satisfactory character handling and peripheral capabilities, did computer use in this field be to make real progress. There is little doubt that the 1400 series has proved to be successful; in fact, y became so dominant that many 360's are still likely to be simulating or emulating them. Some mechanised documentation centres have even chosen to continue using this type of installation rather than to move to later models (ZMD, Frankfurt/Main).

Most of those who have been concerned with computer use in information retrieval in the mid-sixties are likely to have been involved at one time or another with 1401 Autocoder programs. It is an indication of the impact which can be made by a particular piece of hardware or hardware configuration.

It has been mentioned earlier that software rather than hardware may be at present the factor limiting efficient exploitation of our existing computer facilities, i.e. hardware development appears to outrun our capability to apply it to our best advantage - a situation which is perhaps characteristic for much of modern technology and its relation with society.

The speed with which hardware performance is improving has been clearly indicated by Knight[1,2,3]. His analysis evaluated about 250 general purpose computer systems, covering roughly the first two decades of the computer industry, and it showed that as regards cost/performance (number of operations per dollar) computers have been improving by 80-90% per year for commercial computations. If one adds to this the fact that investment in computers increases by 20-30% per year, it suggests that available computer capacity must be roughly doubling each year or increasing 1000-fold in a decade. His analysis also upheld Grosch's law*. Evaluation of future trends in hardware development (Joseph[4]) suggests that the pace is not likely to slacken much during the forthcoming decade. This situation is somewhat dramatically represented in Figure 1 (Ref. 5). One might add that whilst our society is geared to the type of progress indicated by A, it cannot as yet cope adequately with a development such as that indicated by B. This continued rapid development makes it difficult, if not virtually impossible, to decide with some measure of certainty on the hardware configuration economically and operationally best suited to one's requirements for the next 2-3 years, the more so as delivery dates for the latest types of computer hardware, in particular peripheral equipment, tend to be unreliable. As a result of these conditions, the selection of hardware, which aims to meet the criteria of optimum cost/performance, satisfactory delivery date and compatability with existing equipment, tends to become a rather hazardous problem allowing as its usual solution only a compromise of one form or another.

2. DEVELOPMENT AT SHELL RESEARCH LTD., SITTINGBOURNE

The present paper is intended to comment on hardware choice and its problems in the light of the experience gained in the development of mechanised information facilities at the Shell Research Laboratories, Sittingbourne. It is therefore desirable to outline the environment in which we operate, the objectives we have to meet and the resources available etc.

The Sittingbourne Laboratories are a group of Laboratories which deal predominantly with biology-oriented research (Fig. 2), even though perhaps 70% of its research activities fall within the broad disciplines of chemistry (organic and physical) and biochemistry. As a result, much of the technical information activities has been concerned with the handling (storage and retrieval) of chemical information and research data.

---

* Herb Grosch suggested in the late 1940s that computer power increases as the square of cost. The so-called Grosch's law may well be an artifact of the computer manufacturers' pricing policy.

The group dealing with such activities is the Technical Information Services, which are responsible not only for library/information services but also for the provision of digital computer facilities on the site. The group has developed fairly rapidly since its very modest beginnings in 1956 (Fig.3) and is perhaps rather unique in its potential to achieve an effective integration of information service and computer processing facilities. The main impetus for the application of computers in information storage and retrieval arose from the need to create effective information storage and retrieval facilities for our research data.

In 1962 it was decided to move towards a centralised system involving the use of computer facilities. The main stages in this development, which resulted in applying the experience gained in research data handling also in the field of published and report literature handling, are displayed in Figure 4. This figure is a modified version of one given in an earlier paper in 1967 (Ref.6). We appear not to have deviated too much from the course then predicted although, as one might expect, we have slipped roughly a year in the implementation of some projected activities, mainly because we were too optimistic with regard to the impact of time-sharing or, perhaps more correctly, we not only believed too readily the statements made by the computer manufacturers regarding the delivery and cost effectiveness of their systems but also we may have been inclined to take on rather more than we could conveniently cope with.

Also as shown in Figure 5, we have been using a variety of machines since 1962 with its obvious corollary of changes in programming languages. However, it must be conceded that as a result of it all we now appear to have the benefit of a system that, although still developing, has already proved to be very attractive from the operational point of view.

The present system, insofar as it affects us at Sittingbourne, is schematically outlined in Figure 6. Hence we have available on site a small-to-medium size computer installation, the Univac 9300, which can be used as a free-standing system, suitable for file handling and data reduction jobs requiring modest processing power, and it can also operate as a terminal to the powerful Univac 1108 system in Shell Centre, London. Through this we also have access to the 360 system and for transmission purposes to other locations connected to the network.

Figure 7 gives an indication of the type of information retrieval operations which are now handled via our computer system. Some aspects of the system now in operation have been described in various earlier papers[5,7,8,9,10].

Rather than discuss these systems, their method of operation and the reasons for their implementation specifically, it seems appropriate to discuss under a few general headings the experience gained in implementing them.

3. LARGE versus SMALL MACHINES

Mention was made earlier of Grosch's law, i.e. the fact that the capacity of computers appears to relate according to the square of their cost, and as already indicated, this "law" is likely to be an artifact of the pricing policy followed by computer manufacturers (it does seem by no means certain that it will also apply in future).

It has given rise to the dictum "the larger the machine the cheaper it can do a given computing job". This is, as a rule, true if one compares the processor times actually devoted to job execution. However, large powerful processors tend to be input/output bound; this does not seriously affect their performance when dealing with really large computational jobs, i.e. jobs in which very substantial computations are performed on a limited volume of input data, to yield again a limited volume of output data. The situation, however, is quite different if large numbers of different small jobs have to be performed or when the input consists of large files of data, each requiring only a small amount of processing. Earlier expectations that efficient time sharing systems would be able to overcome this drawback have unfortunately not yet been realised. Hence it is not uncommon to find that for small scientific jobs, for data reduction work and for many file handling operations, small systems tend to offer better cost/performance than large machines. This trend has been aided by the fact that a considerable number of small, relatively cheap but quite powerful computer systems have appeared on the market. Attractive, small scientific computing systems can now be obtained in the range £5,000 - £15,000 and systems of the type required for information retrieval activities can be had in the range £30,000 - £75,000. The basic Sittingbourne system, considered as a free-standing unit, would fall in the top end of the latter range.

Although the use of a small-to-medium size system may be economically fully justified for various jobs, in particular those in the information retrieval area, this should not prevent us from clearly recognising the drawbacks. One may have to use Assembler-type languages in order to get the full benefit from the small machine and this may mean re-programming if one should want to run the job on another machine. The fact that one's staff has to acquire this particular programming skill is an additional cost factor, even though time to program a job is probably on average not greater than if a high-level language such as Cobol were used. (In fact, skilled programmers tend to find use of a language such as Cobol rather a bore.) Furthermore, one is not likely to accept lightly a situation where a job, which would have cost 5 minutes on the large machine, is tying up several hours on the small machine. Labour effort (programming, operating) requirement for a given job is perhaps somewhat less favourable for the smaller than for the larger machines (Fig.8); this will be justifiable in those areas where the small computer has an edge over the large one, in others it will not.

All in all then the choice is a complex one and depends on the type of job and the type of access acquired

The safest course and the most attractive one would appear to be the one implemented in our case where one has the choice between running the local system as a free-standing unit where this is clearly warranted and feasible, and using it as a terminal to a large computer system in all other cases.

4. ON-LINE OPERATION versus USE AS FREE-STANDING MACHINE

The foregoing discussion leads us to look more closely at ways and means of using an on-line machine.

In our case, the Univac 9300 was originally conceived as a local data collecting/editing device through which site peripherals would access the Univac 1108 which was expected to operate under a multiprogramming system, the Exec 8. This system would allow the 1108 to be used concurrently for batch processing, for rapid response (demand processing) type of operation as well as for real time operation. Slippage in the delivery of software with the performance characteristics required meant that the 1108 remained operational under batch processing executives Exec II and later Exec III. It also meant that local equipment, such as that installed at Sittingbourne, had to be upgraded, e.g. via provision of 4 tape units instead of 2, in order to enable it to execute jobs on site where this proved desirable for cost or operational (speed of access) reasons.

Operation on-line is, of course, subject to multiple hazards; it is not only affected by failures of the local equipment but also by any mishaps occurring to the GPO lines and the central computer installation. In our case, they appeared more or less equally responsible for loss of jobs or access.

We also found that the 9300 might be occupied for as much as 15-20 hours for every hour 1108 time used. This obviously made for inefficient use of the 9300 or alternatively it added considerably to the cost of 1108 work. In order to remedy this situation, a system modification has been written which will enable the 9300/1108 communication to be handled as a symbiont operation* and this, together with expansion of the 9300 memory by 16 K to 32 K, should make it possible to reduce the time during which the 9300 is solely occupied in communication with the 1108 very drastically indeed. The improvement in 9300 capacity arising from the more efficient communication should just about pay for the extra cost in memory; in addition, of course, the machine has become a more powerful device capable of dealing more efficiently with a larger range of jobs.

One of these jobs, and quite an important one to us, is the operation of our SDI system. This is run on the 9300 at present mainly on the basis of CA Condensates tapes; it has developed into a rather substantial operation covering the use of 150-160 search profiles (ca 1800 search terms) which are used to select each week ca 2000 references (on average about 12 per profile) from the tapes received from Chemical Abstracts Services, Columbus. This facility is now used by the great majority of the research workers in the Sittingbourne Laboratories; it leads, however, to a substantial load (5-6 hrs each week) on the 9300. This load was fully anticipated; in fact, we feel rather proud that the development and performance of the system proved to match our prediction so closely. However, as since then also a variety of other jobs have been implemented on the 9300, not only information storage and retrieval jobs, such as those indicated in Figure 7, but also a variety of jobs concerned with processing of instrument output, one is obviously anxious to reduce such large loads. This will be achieved by utilising to the full the speed improvements potentially available as a result of larger memory. Passing the job onto the larger computer has as yet operational disadvantages (e.g. load on data transmission) but this might eventually well come about. At the same time, there is a tendency, as a result of the availability of very cheap, small computers, to take processing of some of the instrument output away from the site computer and carry it out near the source. Hence, we appear to be moving towards a 3-level system, i.e. processing at source, on the site computer and on the larg[illegible]tral computer system.

The development of this approach, as well as the appropriate volume of usage at each level, will, one expects, be guided by the combined objectives of meeting user requirements, particularly as regards access (turn-around time), and lowest cost for equipment or, in other words, one tends to strive towards maximum productivity from the combined resources of labour and equipment. This tends to guide also the usage of existing equipment. For instance, with regard to the development of new applications, it may be preferable to do this in some cases direct on the large machine, which may have the various routines required in this work already available; in other cases, however, particularly non-numeric jobs, one may well find that development work can most conveniently be done on the smaller computer, to be followed by implementation on the large computer when the job becomes fully operational.

Much depends here on convenience of manipulation and access. In connection with this, it may be of interest to show how access to computer facilities and the volume of computer usage at Sittingbourne have developed during the past six years (Fig. 9). This figure is a modified version of one given in an earlier (1967) paper[4]. The original version reflected the expectation that response time would sharply improve during 1967/68 as a result of the implementation of Exec 8 on the Univac 1108 system. This has not, as yet, materialised and it is of interest to note that the improvement in response time we actually did achieve does fit in much more smoothly with the gradual progress made during 1963-67. Large, complex systems rarely change drastically, they evolve.

---

* A symbiont is an independent routine, which transfers data between a peripheral unit and an intermediate storage medium such as a drum.

In order to display more directly the relation between response time and computer usage, the data in Figure 9 were used to provide the graph plotted in Figure 10. There is, in our view, little doubt that improvement of access acts as a powerful stimulus in computer application.

5. STORAGE FACILITIES

In an account on computer hardware relating to information retrieval work, storage facilities obviously ought to figure prominently. If they do not in the present paper, it is because we have in information, storage and retrieval work little experience to offer beyond the use of magnetic tapes.

Magnetic tapes are obviously the mainstay as regards storage in any computer operated information storage and retrieval system; the storage medium is cheap, i.e. say 10 shillings per million characters stored, and magnetic tape units are now reliable and sophisticated pieces of machinery, allowing the tapes to be scanned at speeds varying from 30,000 cps to 300,000 cps and over. Tape densities have been steadily going up, 1,600 bpi is now being offered on most tapes (equivalent to say 30 million characters per 2,400 ft reel of tape) and it is expected that 10,000 bpi might be a reality by the early seventies (200 million characters per 2,400 ft reel). Hence, there is a lot of life yet in magnetic tape systems and it would seem that it will be hard to beat them on grounds of economy, even with micro-image storage techniques. We may, therefore, assume magnetic tape and magnetic tape units to be around for a long time yet and whereas one may expect the storage capacity of the medium to increase, one will also require greater versatility in magnetic tape units, i.e. to skip and to select specific file segments at high speed. This is likely to become more and more important as capacity of the tapes increases. Such facility in combination with use of discs might provide random access facility at a relatively low cost. As yet, random access, a facility essential for on-line enquiry, can only be provided by storing the file(s) one intends to use on disc or drum. If the files are sizeable ones, this becomes an expensive affair, e.g. £400/year/million characters; hence one tends to load the disc/drum only for selected enquiry or retrospective searching sessions. This obviously does not make for easy/ready access. Whereas the choice is perhaps easy when one has to deal with a relatively small file (up to a few million characters) used frequently because then one can justify keeping it permanently on disc or drum, the difficulty arises with the somewhat larger files which are consulted not all that frequently, and unfortunately such cases are more frequent than the ones in which the choice is a straightforward one. There is no obvious, simple and satisfactory solution to this problem if one wants to tackle it as an on-line operation. In our case, virtually all file consultation/search jobs are handled as tape jobs, carried out according to priority and requiring the tapes concerned to be specially mounted. Hence, response time is at best 15 minutes and at its worst a few hours. The position should get easier as disc storage facilities become available at lower cost but this is not going to make a dramatic improvement in the near future. For some time to come we will have to rely on the judicious use of tape storage and disc/drum storage facilities separately as well as in combination.

Another aspect of on-line storage arises from the fact that increasingly large volumes of machine-readable literature are likely to become available in the near future as a result of the acceptance of computer-aided editing and typesetting in the publishing world. At the same time, it seems more than likely that it will become increasingly difficult for the medium sized libraries to maintain adequate coverage of the relevant literature in view of increasing costs; this may well lead to a rapid increase in the need to borrow literature. It has been suggested[5] that we should consider the establishment of data banks in suitable national centres, e.g. in the UK the National Lending Library, where machine-readable current literature should be collected, to be accessed by local information services via a digital network. Such centres would require very large storage facilities, probably beyond the range of, for example, the Datacell (8×4.10^9 char) but facilities of this type have already been developed experimentally, e.g. IBM's photodigital mass storage system[12]. A more generalised outline of this type of device is given in Figure 11. Such facilities should, one hopes, become more generally available during the first half of the 1970s. Although the storage cost per million characters stored may be expected to be low, they will have to be designed for very large capacities if they are to be economical, hence their total cost is still likely to be high. One can therefore contemplate their use only as part of a large computer system, e.g. our 1108 system or a national system.

Their use again presupposes the availability of a computer network as well as the capability to send and receive locally the information relevant to the operation of such on-line systems.

6. MAXIMISING THE OPTIONS

At several stages in the foregoing discussion, the need for a system allowing various options in dealing with a problem, a search query, has become apparent. This freedom in the selection of a suitable strategy is required not only in order to ensure an efficient match between the processing requirements for a specific problem and the equipment to be used for its solution but also to achieve the most effective match as regards user requirements and system capabilities. As indicated in an earlier paper[13], the manner in which an information storage and retrieval system interacts with its users is perhaps the most crucial factor in determining the economic merit of the information service with respect to the organisation served.

In the design of an information system, one should therefore pay particular attention to the way it allows on the one hand the search strategy to be adapted to the specific query, and on the other, the user to interact with the system.

A fairly good example of this type of approach is our overall system for chemical structure handling[7] with its many options as regards search strategy and user interaction. However, as a meaningful discussion of this system is rather beyond the scope of the present paper, it is preferable to use a simpler case and illustrate the approach with an outline of the manner in which we operate our SDI system. As indicated above (under 4), most of our current awareness requirements on the site (perhaps 70-80%) are now covered by an SDI service which makes use of CA Condensates tapes supplied weekly by the Chemical Abstracts Service, Columbus. It is a system run very much under direct control of the user who can adjust selection for any weekly search by modifying his search profile. The output is retained on tape, a print-out of the hits is supplied to the user and screened by him to indicate articles of interest. This information is used to prepare a user file with relevant papers relating to the profile concerned. To this is also added information obtained via scanning of journals available in the library but not covered by the CA Condensates tape (Fig. 12).

Hence, we have thus created a magnetic tape file which could act as a substitute for the research worker's own private indexing system. In order to make this really acceptable, one has to provide adequate access to this file; at present this is likely to be done via the provision, periodically, of a KWIC index, but at a later stage teletypes will be made available to give the user the means to print-out from his file or add to his file. The economic incentive to this approach arises from the desire to reduce the amount of time the research worker has to devote to current awareness activities and the maintenance of a private indexing system. In this respect, there appears to be a sound economic case for the maintenance of user files and the provision of KWIC indexes. The more direct access via teletypes, however, is as yet more difficult economically as well as equipment-wise.

There is a limit as to the number of peripheral devices one can support with the 9300 configuration available to us, and we are likely to reach it in 1970 when 3 to 4 teletypes will be attached to the system. Further moves will depend on upgrading of the system to Univac 9400 capability.

Access via teletypes is in our case required for:

(a) Data collection (e.g. screening data)

(b) File consultation

(c) Text editing and storage.

In the course of this year and 1970 the implementation of such types of operation is likely to be explored on a limited and experimental basis. Similarly with regard to visual display facilities, we hope in 1970 to explore the feasibility of using this to enable the gradual phasing out of the use and maintenance of the large card indexes of research data which are still being kept operational, despite the fact that the major proportion of all the data is now also stored on magnetic tape. Here again one has to find an economically acceptable solution to providing adequate access without absorbing too much of the overall systems facilities in such an activity.

As a partial solution to this type of problem, we have now installed an automatic feature card punch (Fig. 13). This will enable us to output data held on our magnetic tape files (e.g. compound properties) onto punched feature cards (Fig. 14). Output from the computer is on paper tape which acts as input for the feature card punch. It is thus possible to prepare at a low cost feature card systems tailor-made to specific user requirements. The main application is expected to be in the area of data manipulation and searches, e.g. retrieval according to specific property combinations. It will provide the research worker with a further means of access to data relevant to him, this in addition to the tape searches run on the computer.

7. COMPUTER NETWORK

It may be clear from the above that in our view a satisfactory long term development of information storage and retrieval systems requires the availability of an effective computer network. Fortunately, the Shell organisation in the UK is, for other reasons, already making satisfactory progress in this direction (ref. Figure 6). There would appear to be three main aspects on which one might expect benefits from the network approach.

(i) Actual processing - the feasibility to move the job towards the location/the level at which it can be most effectively and economically executed

(ii) Mass storage - the large capacity storage facilities referred to under 5 can only be justified economically if associated with large centralised installations. Their possible use in relation to the storage of published literature has been discussed in an earlier paper[3].

(iii) Data transmission - one would hope that the network will enable a greater use to be made of direct data transmission replacing at least in part the present procedures based on transport of printed matter (or tapes) and the associated delays.

Developments in this respect are as yet very slow, yet it seems not unrealistic to expect that they will figure prominently in the first half of the 1970s hence, they should be borne in mind even now when selecting equipment.

As an example of the stepwise build up of computer facilities for an information service, leading towards their operation as part of a network, the development projected in Figure 15 (Ref. 11) may be of interest. It concerns the possible implementation of such facilities over a period of about 10 years in a university library environment.

8. CONCLUDING REMARKS

The most significant aspect of computers and their applications is the speed with which this field is developing (ref. Figure 1). The indications are that the impact of this will become very pronounced in the course of the next five years; it should drastically increase our involvement with hardware. The difficulty of choice, however, may well ease as hardware, and particularly peripheral devices, are likely to become more plentiful and much cheaper.

Perhaps many of our yet unfulfilled wishes, e.g. for cheap and versatile consoles, ease of interfacing, on-line mass storage at a reasonable cost, widespread data transmission facilities, will be met in the next few years. We may expect, largely as a result of LSI (large scale integrated circuitry) a significant increase in hardware implementation of software features and in hardware reliability.

We will have hardware in much greater quantity and versatility; the main problem will be how to integrate and use it effectively and economically.

REFERENCES

1. Knight, K.E. — PhD Thesis. Carnegie Institute of Technology, 1965.

2. Knight, K.E. — *Changes in Computer Performance*. Datamation, September 1966, pp. 40-54.

3. Knight, K.E. — *Evolving Computer Performance*. Datamation, January 1968, pp. 31-35.

4. Joseph, E.C. — *Computers: Trends Toward the Future*. Invited papers, I.F.I.P. Congress, Edinburgh, August 1968, pp. 145-157.

5. Dammers, H.F. — *Integrated Information Processing and the Case for a National Network*. Inform. Stor. Retr., Vol. 4, 1968, pp. 113-131.

6. Dammers, H.F. — *Phased Transition from Non-Mechanised Information Storage and Retrieval to On-Line Computer Operation*. Proceedings F.I.D./I.F.I.P. Conference on Mechanised Information Storage, Retrieval and Dissemination, Rome, June 14-17, 1967, pp. 415-436. North-Holland Publishing Company, Amsterdam, 1968.

7. Dammers, H.F., Polton, D.J. — *Use of the IUPAC Notation in Computer Processing of Information on Chemical Structures*. J. Chem. Doc., Vol. 8, 1968, pp. 150-160.

8. Polton, D.J. — *Conversion of the IUPAC Notation into a Form for Computer Processing*. Inform. Stor. Retr., Vol. 5, 1969, pp. 7-25.

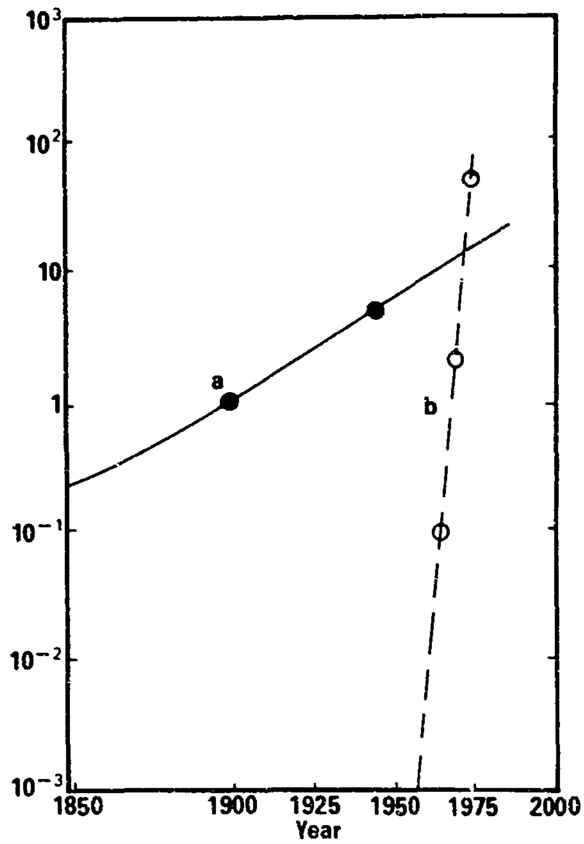
9. Gallagher, P.J. — *Some Experiences with Selective Dissemination of Information*. Information Scientist, Vol. 2, 1968, pp. 103-106.

10. Dammers, H.F., Gallagher, P.J. — *Computer Use in Information and Data Handling*. Automation, Mechanisation and Data Handling in Microbiology, Society for Applied Bacteriology Technical Series No. 4, Academic Press, 1969. (In press)

11. Dammers, H.F. — *Computer Systems in Academic Libraries - Some Aspects of Planning and Evaluation*. Proceedings of the University of Newcastle Seminar on The Management of Computing Activity in Academic Libraries, Oriel Press, Newcastle, England, 1969.

12. Kuckler, J.D., Kerby, H.R. — *A Photo-Digital Mass Storage System*. AFIPS Conference Proceedings, Vol. 29, Spartan Books, Washington, D.C., USA, 1966, pp. 735-742.

13. Dammers, H.F. — *Computer Use in Information and Data Handling: An Appraisal of its Economic Aspects*. Paper presented at the Second International Conference on Mechanised Information Storage and Retrieval Systems, Cranfield, 2-5 September 1969.



Growth of mech. energy relative to human muscle power (a)
Growth of computer capacity relative to human resources (b)
For clerical work (acc. to Ref. 5)

Figure 1

Laboratory	Subject field
Woodstock Agricultural Research Centre	Chemicals, in particular pesticides, for use in agricultural and public health
Tunstall Laboratory	Toxicology, environmental health aspects of Shell products/processes
Milstead Laboratory	Chemical enzymology; natural products of biological significance; microbiological research

Total staff 500 +

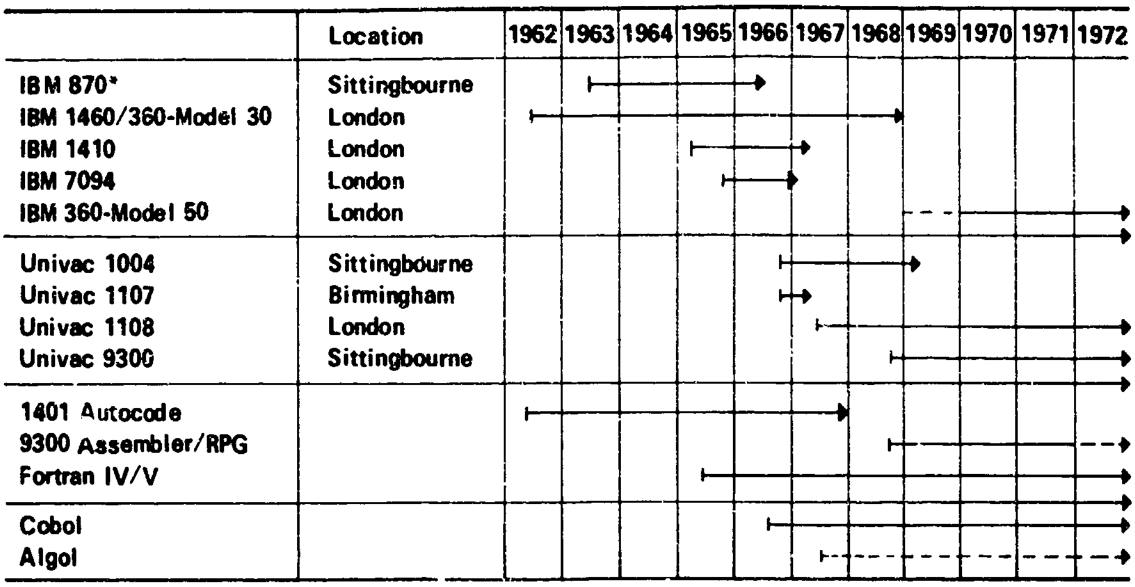
Fig. 2 Shell Research Laboratories at Sittingbourne, Kent

Stage		Prominent staff activity type	Major type of activity
I	1956-1962	Technical librarian	Building up of library (holdings) Building up of subject filing system
II	1962-1968	Information specialist	Literature searches/surveys Research data coding and storage Current awareness Development of computer application in ISRD
III	1968-	Information technologist	Development and maintenance of computer operated information systems allowing maximum user interaction (SDI, research data dtorage and retrieval)

Fig. 3 Technical Information Services 1956 - Shell Research, Sittingbourne

	Main developments
1962 - 1964	Preparatory phase Start with research data coding (structures) IBM 870 tabulation IBM 1401 and Autocoder use Feature card system development
1964 - 1966	First computer operated search system operational (chemical structures) KWIC indexing Variety of computer systems (processors) used
1966 - 1968	On-line computer use operational SDI - Data acquisition on magnetic tape Tape typewriters for text and graphic structure input
1968 - 1970	Establishment of various access points on site Fully mechanised feature card input Start of computer aided typing Start of real on-line mass storage Rapid response query systems Use of visual display
1970 - 1972	Multi-access system widely extended (perhaps 20-30 access points) I/O improvements (CRT, Rand tablets etc.) Capture of most site-produced texts for computer editing/storage Sizeable on-line mass storage of literature information Gradual disappearance of most manual retrieval tools

Fig. 4 Developments in mechanised information storage, retrieval and dissemination at Shell Research Limited, Sittingbourne



* The IBM 870, though not a computer, has been included as a forerunner to site computing facilities

– – – – – – indicates minor use only

Fig. 5 Computer facilities/languages

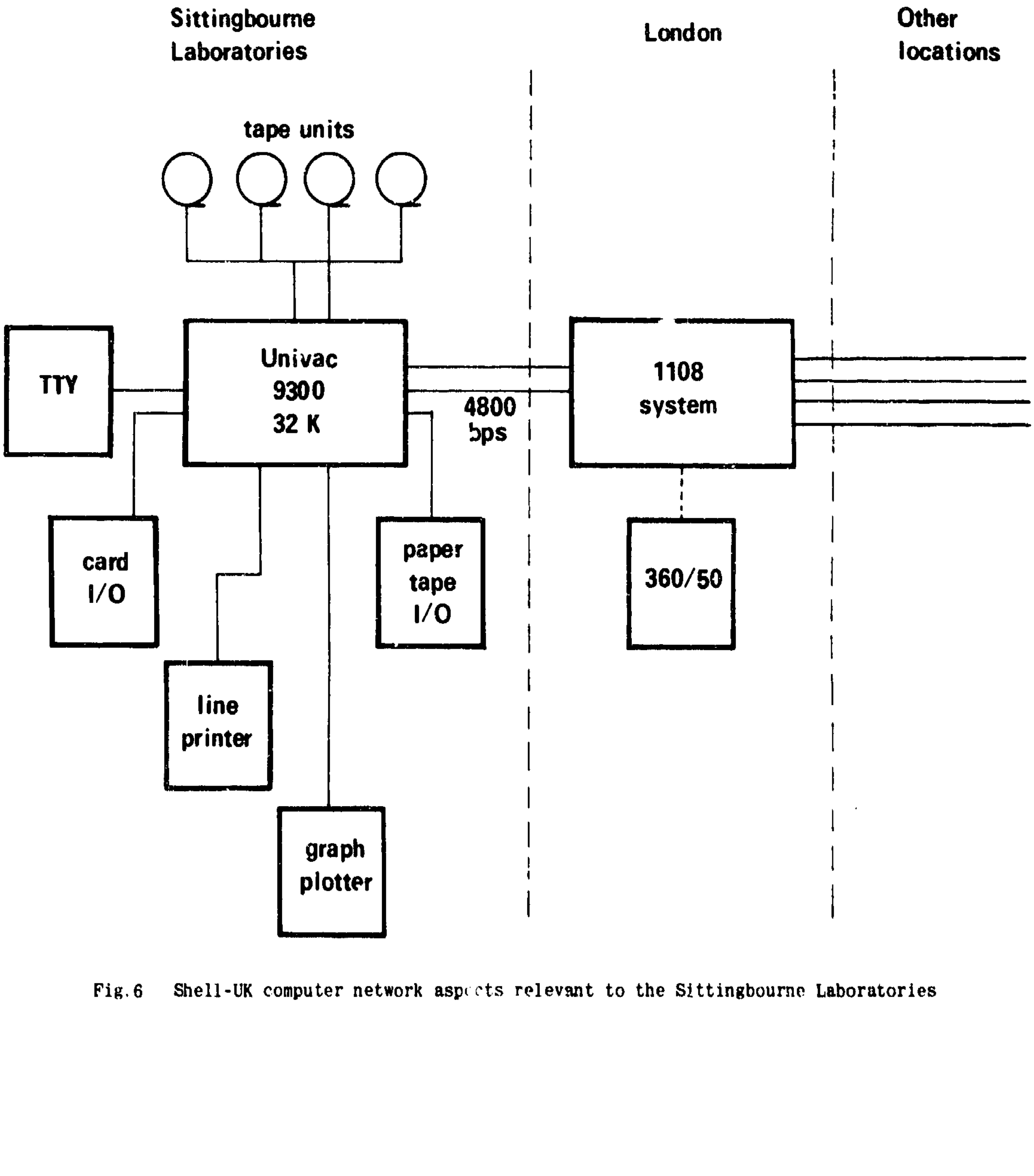


Fig. 6 Shell-UK computer network aspects relevant to the Sittingbourne Laboratories

RESEARCH DATA

Chemical structure storage and retrieval (50,000 compounds)
Storage and retrieval of biological screening data
Storage and retrieval of information on spectra

COMPANY REPORTS LITERATURE

KWIC indexes
Tape search according to profile in preparation

LITERATURE INFORMATION

SDI - CA Condensates tapes, ca 150 profiles, 1800 terms, weekly
Bibliographies
KWIC indexes
Tape searches of patent literature (Derwent records)

LIBRARY OPERATIONS

Control of periodical holdings and accounts
Loan control
Meetings lists

Fig.7 Computer operated information retrieval activities at Shell Research, Sittingbourne

	Univac 9300	Univac 1108
Performance ratio for many 'compute bound' jobs	1	50 - 70
Performance ratio for many '1/0 bound' (IR type) jobs	1	10 - 20
Performance ratio as regards capacity related to labour effort (Programming, operating) required	1	25
Hardware cost ratio	1	25

It should be stressed that the data in the above tabulation are very tentative only and should be looked upon as merely the relative order of magnitude.

Fig.8 Performance ratio for large (Univac 1108) and small (Univac 9300) computer systems

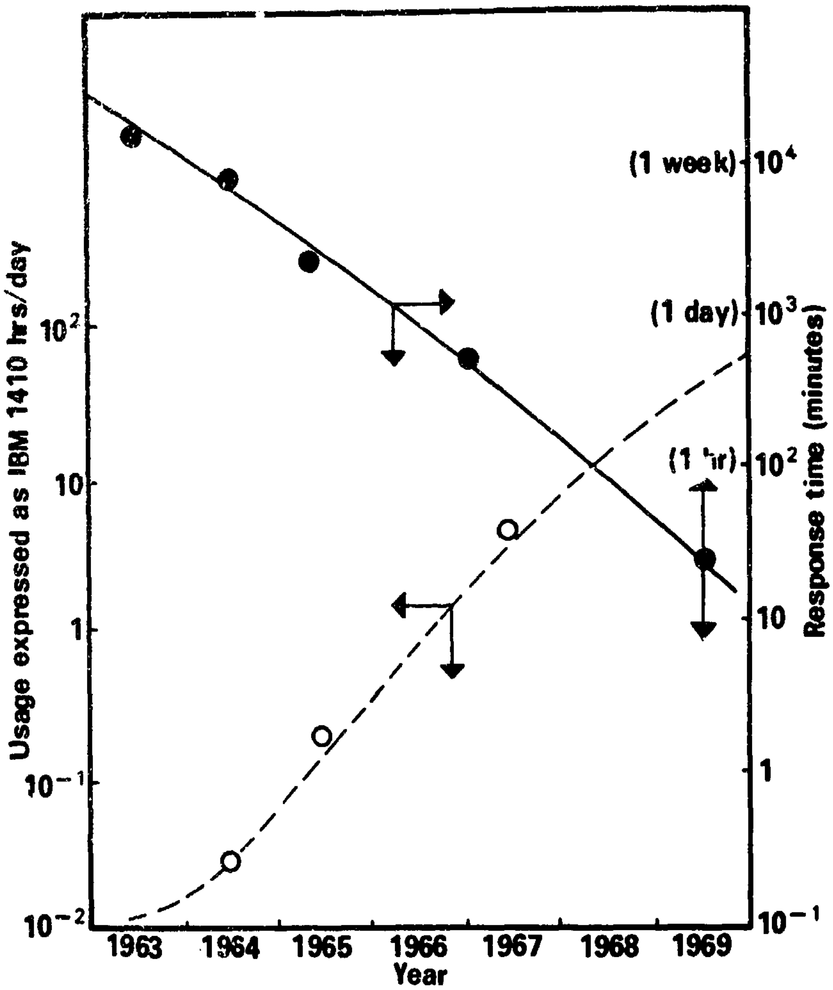


Fig. 9 Sittingbourne Laboratories computer usage/response time

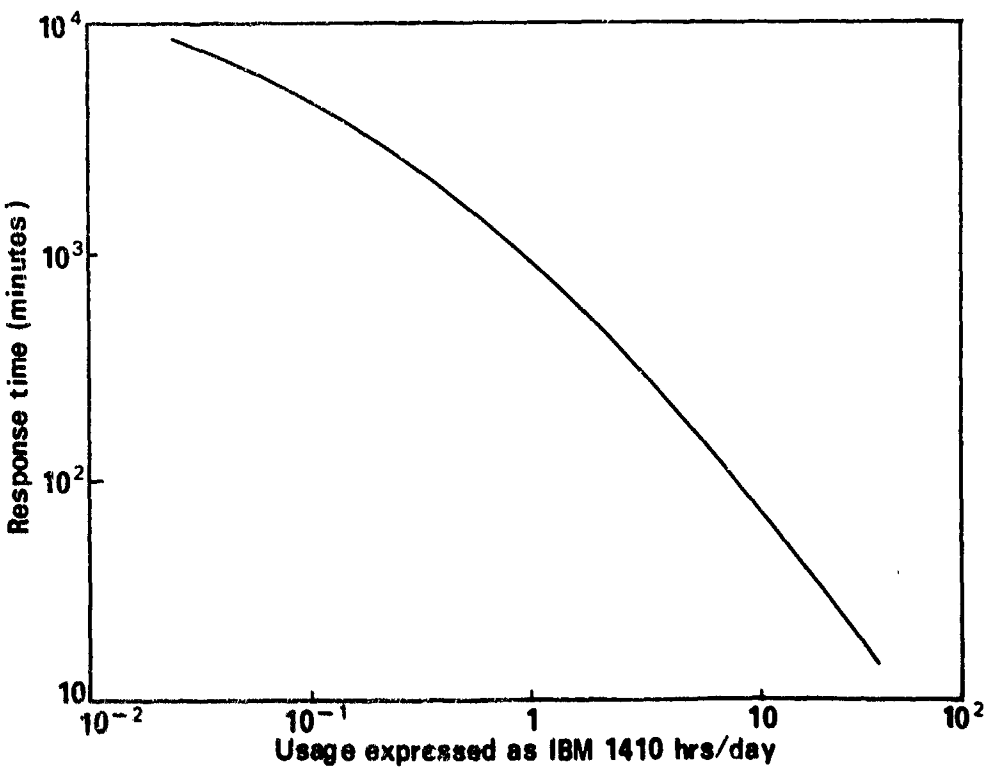


Fig. 10 Sittingbourne Laboratories computer usage/response time

1 Capacity say 10^{12} bits = equivalent to ca 100,000 vols. print.

2 Would appear entirely feasible with present day technology; various design configurations possible.

3 Storage costs will be low, say $1 or less per million bits stored but only for very large capacities.

4 Hence only very large units will be economical, cost say minimum $500,000.

5 Likely to be mixed electronic/mechanical. One cannot visualise low-cost electronic devices for this purpose, at least not in near future.

6 Access to random file location will be relatively slow, say of the order of a second but with suitable dynamic multi-level storage organisation this should be no objection for ISRD use even in real-time operation.

7 Read speeds should be quite high, say ¼ to ½ million characters per sec.

8 Likely to be read-only stores, but again for archival purposes this should be quite acceptable.

9 Available experimentally 1967; perhaps in more general use by early 1970's.

Fig. 11 Mass storage device

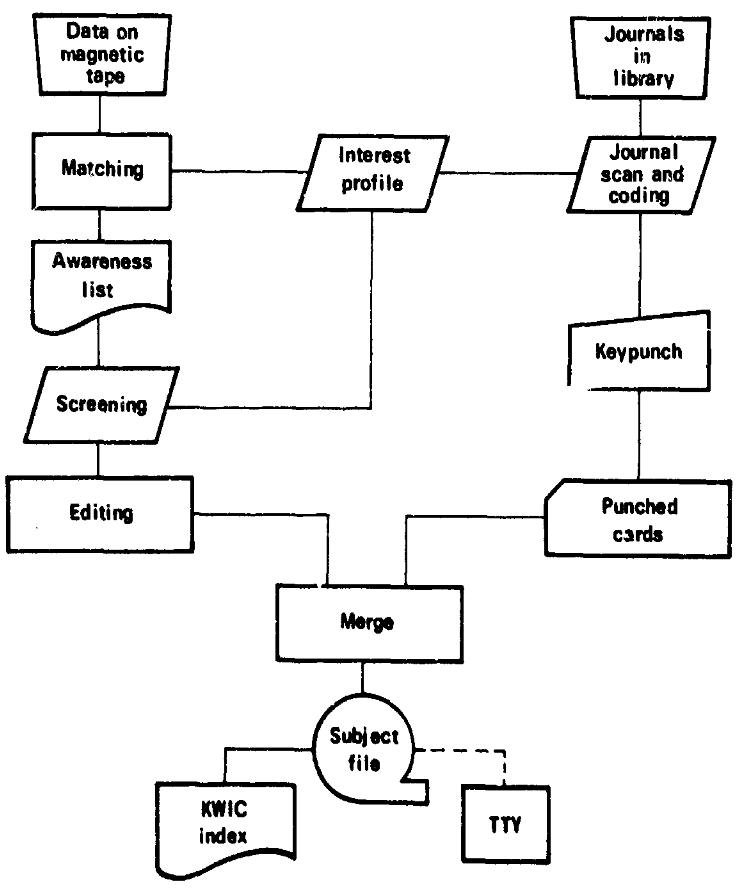


Fig. 12 Computer handling of published literature using magnetic tape and printed journal input

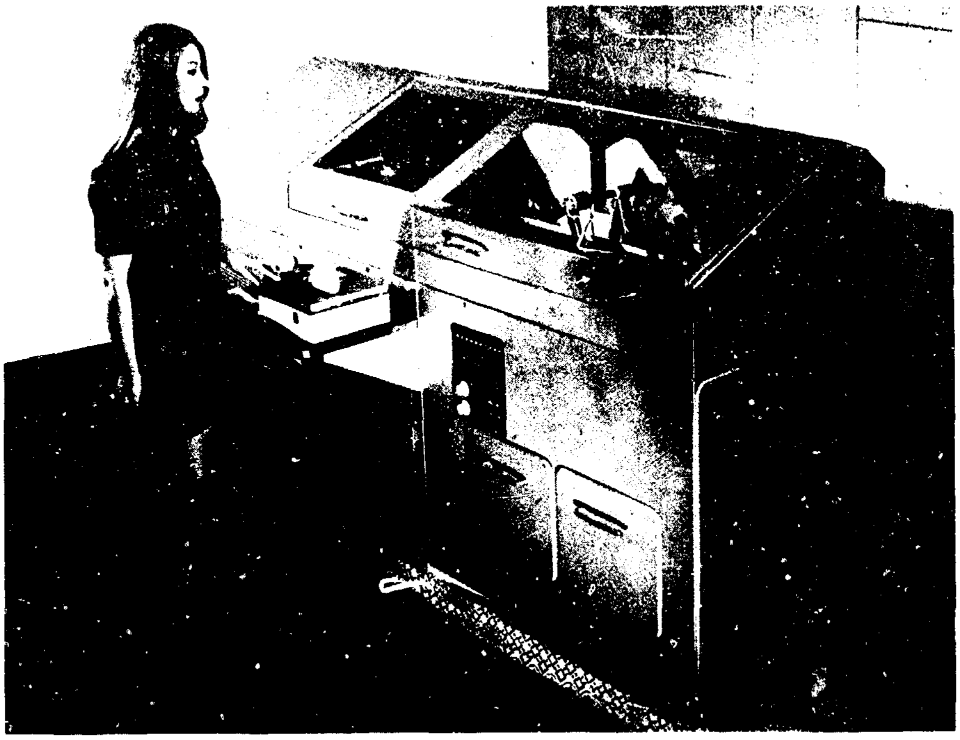


Fig. 13 Automatic feature card punch

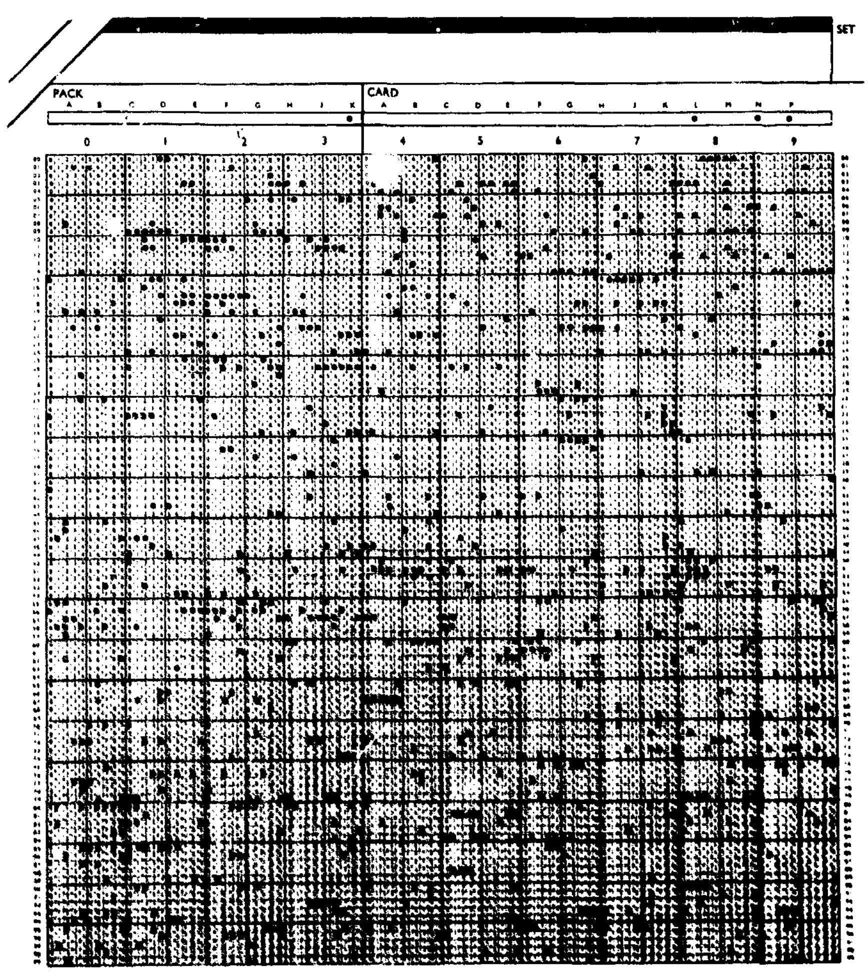


Fig. 14 Punched feature card produced via the automatic feature card punch

Phase	Computer budget	Equipment	Applications
I	ca £3,500	Keypunches; tape typewriters; off-line computer use	Loans system, journal listing, journal subscriptions, bibliographies
II	ca £13,000	DCT 2000 type terminal + card and paper tape read/punch; on-line processing	Building up of files in central computer
III	ca £27,000	Library computer on-line to large university computer; Several magnetic tape units	Processing of magnetic tape files in library. Exchange of tapes
IV	ca £46,000	Several access points to library computer. Tapes and discs, multiprogramming capability	Establishing of on-line enquiry systems and display
V	ca £72,500	Graphic displays. On-line query	Operation as part of large on-line network

Fig. 15 Computer facilities in university libraries: phased implementation

PAPER 7

THE INFLUENCE OF SYSTEM DESIGN ON SYSTEM QUALITY AND ECONOMY

by

Rudolf Bree

Director of the Technical and Scientific Documentation Centre
of the European Community, Luxembourg

SUMMARY

System's design for mechanized information services must try and aim at a reasonable balance between the quality of the service offered and the operational economy in rendering services. Considerations on input operations, subject control, structurization of the vocabulary used in the system and on output-forms are discussed in respect to their effect on quality and operational economy. The experience gained during the development and the operation of the nuclear documentation system of the European Community served as basis for the presentation.

SOMMAIRE

Dans la conception d'un système destiné à des services d'informations mécanisés, on doit s'efforcer de réaliser un équilibre raisonnable entre la qualité du service offert et l'économie opérationnelle liée à la réalisation de ce service. L'auteur étudie les opérations d'entrée, le contrôle des sujets, la structuration du vocabulaire utilisé dans le système, et les formes de sortie, à la lumière de leur influence sur la qualité et l'économie opérationnelle.

L'auteur s'est basé, pour la préparation de son exposé, sur l'expérience acquise au cours de la réalisation et du fonctionnement du système de documentation nucléaire de la Communauté Européenne.

THE INFLUENCE OF SYSTEM DESIGN ON SYSTEM QUALITY AND ECONOMY

Rudolf Bree

Introductory Remarks

The topic, on which I have been asked to speak - that of system design with reference to small documentation centers - makes me feel slightly uncomfortable, for the simple reason that I don't think that small documentation centers will have to face such problems.

Smaller documentation centers will normally have very little choice in this context - if the category "small" documentation center is used in the sense of a limited staff serving a limited number of clients.

This leads to the conclusion that small centers of this kind must depend upon larger ones.

Consequently "system design" is something they might never have to cope with and this prevents them rather effectively from committing errors in this field. They can leave this happily to the larger centers. However, cooperating with larger centers they are tied to the mistakes these commit.

Therefore, system design must be of interest to smaller centers. They should understand the conditions upon which certain decisions have been based. Being a smaller center means also being much closer to the individual user. This puts them right between these users and the system designers.

I could stop here by stating that anything more said is beside the point. But it might still be interesting to discuss errors in system design, regardless of the size of the center. My remarks might be biased. They are based upon specific experience with operation of the EURATOM system for storage and retrieval of nuclear information, which was developed rather early on and without much help from experience of third parties. Therefore ample room was left for the method of trial and error.

Which main features did we have in mind when we mapped out that "system"? We wished to assist the potential client to solve his particular problem by giving him access to a maximum of the available information. What he was looking for he should get as fast as possible, at least before his mind got stale on that problem and before he decided to abandon his search for available information and start experimenting all on his own, whatever the expenditure in money and time might be.

Furthermore we wished to supply information to the client in a form and a quantity he could digest, and in a way which would help him without burdening him. So, what we were striving for in our system design was both useful and usable quality - within the terms of a justifiable budget for development as well as for operational economy. We felt that without quality we would lose the users and without operational economy we would risk losing our sponsors.

The average user does not seem to be interested either in the degree of sophistication of the system or in the methods by which operational economy is obtained. What he cares about is being served in a speedy and reliable way and in a form which makes it easy for him to digest the offered material.

Trying to meet these simple sounding requirements in system design is not so easy. A few of the design aspects we can speak about from our own experience are:

1. Input based on titles, abstracts or complete texts;
2. Subject control by free or by controlled vocabulary - or by a hierarchical system;
3. Structurization of a controlled vocabulary: is it possible and is it helpful?
4. Should the retrieval organization be based upon Boolean algebra or on weighting?
5. Output in the form of title listings, of abstracts or of complete texts?

1. INPUT BASED ON TITLES, ABSTRACTS OR COMPLETE TEXTS

Quality of retrieval and operational economy are both greatly influenced by the form of input chosen.

1.1 For access to a document, knowledge of its *title* seems to be indispensable. Basing *input* solely on titles is impossible because they are generally insufficient for subject representation, and therefore for subject control. It might be necessary to store the titles in the context of the complete descriptive cataloguing of the document.

1.2 *Abstracts* are superior to titles, insofar as their quality is satisfactory. We know that in our field about 8% of the abstracts are badly written and that a further 7% of them could bear improvement. As a basis for retrieval a kind of abstract is required which refers to all the important concepts contained in a document.

Using abstracts for input does not necessarily mean that their texts have to be incorporated in the storage. The limitation of the number of acceptable characters and diacritical signs would still, in most of the existing computers, require a considerable amount of editing of the text, which can be very expensive, especially if texts in several languages are to be used.

1.3 Full document text storage in digital form does not seem to be an economical proposition at the present time. It is regrettable that computer-controlled image storage is not yet in existence. Compared with indexing from abstracts, indexing from complete texts has both advantages and disadvantages. It would compensate for the about 15% bad abstracts. On the other hand, indexing from complete texts alone is time consuming and invites incorporation of too many peripheral aspects of a document in the indexing. This would result in retrieval of non-significant documents. As a general rule, a workable compromise seems to consist in basing indexing on abstracts, keeping the documents at hand for reference and to supplement the indexing if necessary.

It would be an error to base indexing on titles alone, because of the resulting low quality of such indexing.

2. SUBJECT CONTROL BY FREE OR BY CONTROLLED VOCABULARY

Use of a controlled vocabulary - the thesaurus approach - necessitates initially a great many decisions on the selection of the terms. Later on during indexing and retrieval it requires intimate knowledge of the chosen concept. The terms of the controlled vocabulary look like terms of the free, that is the natural, language - which is deceptive. In fact, in some cases many of the terms used cover only part of what the term implies in natural language. Hence the need for profound knowledge of the vocabulary and consistent application during indexing and retrieval.

In these circumstances the idea of doing away with controlled vocabulary and instead using freely the significant terms as used by the author of a paper and readily available in the text is enticing. Well, in practice a "free vocabulary" is not so free. This is because one has to apply much sophistication in storing the terms to avoid an enormous fixed storage in the computer and to facilitate the matching process during the retrieval operation. To take care of the different grammatical forms of each term, truncation can be used. Homonyms and synonyms create tricky problems. Practically all of the numerous terms and forms of the free language have to be reduced again inside the computer to a smaller number of terms - we're back to a controlled vocabulary again. But at the free language end the computer must accept an almost unlimited number of terms which must be made recognizable. Considering that the average author's discipline in applying his terminology leaves something to be desired and that things are complicated further when several languages are to be used, one might soon find that the use of free language is costly and hardly improves precision of retrieval.

It is true that indexing based upon a controlled vocabulary requires a high degree of discernment, but for each document only once for all. Precise focusing in retrieval then becomes possible. So it might be preferable to use a controlled vocabulary, on the condition that certain rules are respected during its development: synonyms and homonyms must be discarded; no term can have more than one meaning.

During the development of a controlled vocabulary, it should be kept in mind that its purpose is not so much to facilitate indexing, but to effect precision in retrieval. However, in the initial phase retrieval usually cannot be tested in a conclusive way. During this part of the development it is recommended to follow closely the frequency of assignment of each term. Too frequently used terms, especially those of too general a character, lose any significance for retrieval because of their frequency. They must therefore be replaced by more specific terms. The EURATOM Thesaurus introduced therefore a category of terms designated "forbidden terms" - which are listed, but only with reference to other terms.

Hierarchical systems have always been dear to many people. Their structure, built around the generic context of terms, seems to offer a very acceptable means of putting order in conceptions and ideas. We cannot judge from broader experience whether or not there is a real chance left for profitable use of such systems in mechanised storage and retrieval systems.

Compared to the two other methods, coordinate indexing has established its feasibility even in large systems - for retrieval quality as well as for operational economy. For smaller centers with specialised collections it is even relatively easily transferred to manual retrieval, using peek-a-boo cards.

3. STRUCTURIZATION WITHIN A CONTROLLED VOCABULARY

A system of coordinate indexing based upon controlled vocabulary might seem to be incompatible with structurization. We found, however, that this can be even combined in a very attractive way. The machine dictionary containing all the thesaurus terms can be used for generic term grouping.

When the EURATOM Thesaurus was compiled, we used the subject headings from NSA and arranged them - in the beginning, to become clear in our minds as to their semantic context and their scope - in the form of two-dimensional graphs. Then we transformed the subject headings into univocal keywords, eliminating all homonyms and synonyms. Maintaining the initial graphic display, we went one step further and arranged in clusters, around the keywords, terms of higher specificity. In the end we had charted the whole nuclear field, depicting visually the relation between the terms, delimiting the meanings of the keywords of higher generic character by the more specific terms around them. We avoided in this way having to compile a glossary and soon we learned how profitable and helpful this kind of display is for effecting sufficient consistency in decentralised indexing. Later on, for formulating retrieval questions, we found it as useful as for indexing, especially after we indicated for each term the frequency of its assignment, because this enabled us to calculate the probable number of answers to be expected during retrieval.

This visual display has great appeal. Supplementing, as it does, the indispensable alphabetical listing of terms, it is in many instances superior as a working tool.

Later on we introduced this generic structure into the machine system itself, starting very successfully an automatic generic posting operation. This permitted to establish the very simple indexing rule: "use the most specific index term". The computer adds automatically all the generic terms. In this way we virtually double the number of terms assigned, facilitating retrieval at any generic level.

In our opinion this kind of structurization and the automatic generic posting would facilitate the work of smaller information centers too.

4. USE OF BOOLEAN ALGEBRA, OR WEIGHTING

The application of Boolean algebra to retrieval in a system based on coordinate indexing does not raise many problems. Its efficiency has been proved. Weighting of terms is an alternative to Boolean operators in retrieval*. In some cases a query is more easily formulated in Boolean algebra and in others by weighting. The selective effect of Boolean algebra is considerable. In a system which embraces a million entries, three keywords, each of which has been used about ten thousand times, can be combined in Boolean algebra in such a way as to produce only about ten references.

However, we have learned in the meantime to combine very usefully both methods, securing firstly satisfactory recall by way of a Boolean term combination and subjecting the result to a second machine routine, then applying weighting, achieving in this way a ranking procedure which helps to eliminate noise.

5. OUTPUT IN THE FORM OF TITLE-BIBLIOGRAPHIES, OF ABSTRACTS OR OF COMPLETE TEXTS

Decisions on output form influence organization and operational economy considerably. Here we have the psychologically important interface between the system and the individual. The more or less elaborate presentation of document data can be justified finally only in terms of its actual usefulness to its clients. The user must feel that he is being helped and not burdened, when using the system.

This has many consequences, most of which seem to be so obvious that it hardly seems worth mentioning them.

One absolute condition of a system is that it should be adaptable to the purposes of the individual user. The number of references which will be acceptable varies greatly, depending on whether the user is writing a thesis, or a handbook, or on whether he is posing his question with a view to applying for a patent, or again on whether he wants to bring himself up to date on recent developments in a given field. We can take it for granted that this adaptation can be effected during retrieval.

So, how can the client best be served: by a title list, a set of abstracts or by complete document texts?

The quickest (and not too expensive) way would seem to be a machine printout of bibliographical data - supplemented, on demand, by the printout of the indexing terms. However, the user cannot read at once all the documents referred to him nor does he usually want to. He wants to start with the most promising ones. How can he select them most easily?

A machine printout of bibliographical data will permit such selection only rarely, so he would have either to select at random or to pass the total list to his librarian asking for all documents listed - neither method being really satisfactory.

Passing on to the user the results in the form of abstracts would obviously facilitate selection and ranking most effectively. However, this is a method which seems to be rather expensive when performed by the computer, not only because of the considerable tape length which results from storing text in digital form, but because a large proportion of the abstracts has to be edited carefully before input in view of printout limitations. As

* Weighting of retrieval terms in order to obtain an output ranked by probable relevancy must not be confused with the (seldom used) weighting of index terms, the purpose of which is to show the relative importance of individual concepts in a document.

an alternative the printed abstracts can be stored, in reproduced form if necessary, in a large file. This is fine for readability. However from the point of view of operational economy it leaves much to be desired. The organization of such a file supposes the availability of printed abstracts. Improvements might be possible by transfer to microforms, but really satisfactory methods of fast access to large files are lacking, even for microforms.

To furnish the user full texts - which obviously would be possible either in full size, reduced size or microcopy, would be even more expensive and would suppose a considerable organization. Moreover, a large part of the effort involved in doing so would be evidently wasted, because documents retrieved by the system but discarded by the user would have been produced unnecessarily.

Consequently the relatively most useful form for output seems to be abstracts, because it really allows the user the indispensable facility to choose between the documents. The higher cost, compared to title printout, is balanced by the gains in user ease and that is, finally, in user acceptance of the service offered.

Where input and output are concerned numerous decisions must be made, which can have far reaching consequences, either for the quality of retrieval, operational economy or the usefulness of the service to the user. I have only mentioned a few as a reminder of the complexity of the choices facing the system designer. To go back on earlier decisions on the form of input or output is always possible at a later stage and would only involve material losses. Errors in design of the subject control could have grave consequences from which smaller centers would suffer as much as larger ones.

PAPER 8

PROBLEMS OF INPUTTING DATA FROM OTHER SYSTEMS

by

Jack E. Brown

National Science Library of Canada, Ottawa

PROBLEMS OF INPUTTING DATA FROM OTHER SYSTEMS

Jack E. Brown

The National Science Library of Canada serves as the focal point of a national scientific and technical information network. Working in close co-operation with other libraries and information centres, it serves as an information transferral agency and endeavours to provide scientists, engineers and industrialists with immediate access to publications and information required in their day to day work.

The NSL employs a variety of computerized techniques to facilitate the handling of information and has since 1966 operated an SDI program in the Ottawa area. In January 1969 the program was redesigned and expanded to provide a National SDI service which has been fully operational since February 1969. A variety of data bases, then on the market, were tested and evaluated to determine the minimum number of tape services which might be exploited to obtain maximum coverage of the S & T literature.

In order to overcome the problems created by incompatibility of tape services, a standard file format, based on the MARC II format was developed, to which various data bases can be converted with relative ease. To date 220 librarians, scientists and information specialists, from all parts of Canada, have been trained in weekly seminars conducted by the NSL, to serve as search editors and provide direct assistance in setting up interest profiles. A "Profile Design Manual" was prepared and distributed to search editors and to all subscribers to the service.

At present, subscribers to the CAN/SDI service have access to three tape services: Chemical Titles, Chemical Abstract Condensates and ISI Source tapes. INSPEC tapes and others will be added as time, staff and funds permit. All journals covered by these tapes are held by the NSL and photocopies of cited papers can be obtained from the NSL, if the journal is not available locally. Subscribers receive computer print-out of pertinent citations, weekly or every two weeks, depending on the tape service selected. References are printed separately on a two-part slip, one part of which is used as a response form to provide feedback and permit updating and improving of interest profiles.

Note: A detailed account of the CAN/SDI project will appear in - "National Science Library of Canada SDI Program" *Special Libraries*, Vol.60, Oct. 1969.

PROFILE - NUMBER	INSERT YOUR ADDRESS LABEL IN THIS BLOCK
142	Dr. S.G. Calvert Division of Biosciences Room 3143 100 Sussex Drive Ottawa, Ontario 142
SHEET - NUMBER	
1	

STATE YOUR SEARCH REQUEST IN NARRATIVE FORM. ADD TWO REFERENCES OF PAPERS PUBLISHED BY YOU OR A COLLEAGUE WORKING IN YOUR FIELD. (PLEASE TYPE OR PRINT)

In order to support ongoing research, I request references dealing with the effect of 2-deoxy-glucose and 2-deoxy-galactose on animals or humans - especially, has anyone administered either of the above and what happened. I am interested in their possible action on an antibiotic, and their toxicity, particularly on humans. The 2-deoxy-glucose is usually cited as 2-deoxy-D-glucose or -d-glucose, and both are also spelled as 'desoxy'. The galactose form has been cited also as fucose. I am interested in the isomers. 3-deoxy-glucose and 3-deoxy-galactose.

REFERENCES

1) J Brown
Effects of 2-deoxy-glucose on carbohydrate metabolism. Review of the literature and studies in the rat.
Metabolism Vol.11, p.1098-1112, Oct. 1962.

2) C Lutwak-Mann
Effect of 2-deoxy-D-glucose on the rabbit blastocyst.
J. Reprod. Fertil. 10, 133-5, Aug. 1965.

LIST PROFILE WORDS AND SEARCH EXPRESSIONS ON REVERSE SIDE

Fig. 1(a) A search profile

TT	W	AC	PROFILE WORDS	TT	W	AC	PROFILE WORDS
T	9	A	2 DEOXY GLUCOSE				
T	9	B	2 DEOXY GALACTOSE				
T	9	C	2 DESOXY GLUCOSE				
T	9	L	2 DESOXY GALACTOSE				
T	9	E	2 DESOXYGLUCOSE				
T	9	F	2 DESOXYGALACTOSE				
T	9	G	2 DEOXYGLUCOSE				
T	9	H	2 DEOXYGALACTOSE				
T	9	I	2 DEOXY D GLUCOSE				
T	9	J	2 DESOXY D GLUCOSE				
T	9	K	FUCOSE				
T	9	L	3 DEOXY GLUCOSE				
T	9	M	3 DEOXY GALACTOSE				
T	9	N	3 DESOXY GLUCOSE				
T	9	O	3 DESOXY GALACTOSE				
T	9	F	3 DESOXYGLUCOSE				
T	9	Q	3 DESOXYGALACTOSE				
T	9	R	3 DEOXYGLUCOSE				
T	9	S	3 DEOXYGALACTOSE				
T	9	T	3 DEOXY D GLUCOSE				
T	9	U	3 DESOXY D GLUCOSE				
T	1	V	ANTIBIOTIC*				
T	1	W	RAT*				
T	1	X	MOUSE				
T	1	Y	MICE				
T	1	Z	ANIMAL*				
T	2	AA	TOXICITY				
T	2	AB	HUMAN*				
A		AC	BROWN J				
A		AD	LUTWAK MANN C				

EN	S	NM	TW	SEARCH EXPRESSIONS
1	C	10	9	(A|B→AB)
2	C	2		AC|AD

Fig. 1(b) Profile words and search expressions (on reverse of search profile)

INFRARED or ABSOR and PROPERTIES or CONSTANTS or REFLECT or REFRACT or SPECTR

THIN FILM and INFRARED or OPTICAL or REFLECT or TRANSMI

INTENSIT and DISPERSION or INFRARED or REFLECT or REFRACT or VIBRATION and ABSOLUTE or ABSOR or CONDENSED or FILM or INTEGRATED or LIQUID or OPTICAL

FOURIER TRANSFORM or CORRELATION FUNCTION or MOLECULAR or ROTATIONAL and MOTION or RELAXATION and INFRARED or MOMENT or RAMAN or SPECTR

SPECTR and CHANNEL or SHAPE and LINE or BAND

ATTENUATED TOTAL REFLECT or ATR or INTERNAL REFLECT and SPECTR or INFRARED

Fig. 2 Example of how a search request may be documented by a researcher

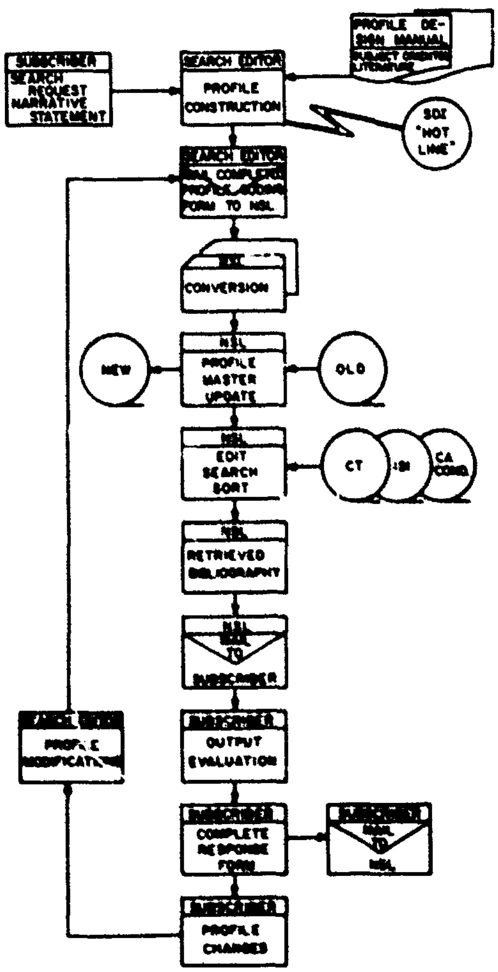


Fig. 3 CAN/SDI project: generalized flow chart

PAPER 9

RETRAINING OF PERSONNEL TO FIT INTO AUTOMATED SYSTEM

by

J.H.Klopp

Cedocar, Paris, France

SUMMARY

Criteria to be considered before introducing computerized information retrieval service(s) are analyzed. Special attention is devoted to decision taking in switching to an automated system and to the training of the personnel including directors and senior staff.

SOMMAIRE

L'auteur analyse les critères qu'il importe d'examiner avant d'introduire un ordinateur dans un service de recherche d'informations. Il consacre une attention toute particulière à la décision de passer à un système automatisé, et à la formation du personnel, y compris le directeur et les cadres supérieurs.

RETRAINING OF PERSONNEL TO FIT INTO AUTOMATED SYSTEM

J.H.Klopp

This is an extremely broad field, and I am afraid my report will only summarize a few evident facts which you all know.

To deal with this subject, I have adopted the following layout:

- Criteria to be considered before you decide whether or not to introduce a computer in your service.
- Use of existing personnel in the light of the aims pursued.
- Re-training.

If you have decided to use a computer, this means that, after seriously analyzing a certain number of factors, you have decided to create a new documentation research system.

1. Let us review those criteria which, in fact, indirectly define your work and the personnel to use:

1.1 Area covered by the system, with indication of the degree of depth for each branch of this area, and of the period covered since the origin.

1.2 Review of services covered by the system, with indication of the extent of the field, of the degree of complexity or improvement reached in the research process, and list of ... points of potential customers.

1.3 Possibilities of extending the system to other fields, with indication of possible ease of coordination with similar services already ensured in initially covered fields.

1.4 Quickness of system, with indication of the speed at which questions on a given document can be answered, reckoning from the date of publication of this document.

1.5 Cost of exploitation of the system with study of cost of:

- data recording
- data processing and output
- use of data provided by the system.

Apart from the factors listed above, other technical criteria have to be considered, such as:

(1) *Card-index size factor*

- number of articles in the card-index
- acquisition rate (increase)
- obsolescence rate
- complexity of field covered

(2) *Activity factor*

- frequency of questions per day, per user, etc.
- necessary speed of answer

(3) *Penetration Factor*

- depth of penetration into given field (indexing depth) of documents included in the information system
- implementation of machine possibilities sufficient for ensuring the above penetration

(4) *Control factor*

(5) *Quality factor*

- reliability of document analysis
- reliability of documentary research operations and, consequently, noise level

(6) *Expansibility factor*

- possibility of adding new notions to the system
- ability of the system to adapt itself to topics dealing with other fields
- possibility of bringing up to date many copies of the overall card-index in various places, at a reasonable cost
- possibility of dealing with a great variety of questions at a reasonable cost

(7) *Cost factors*

- cost of analysis
- cost of codification
- cost of research
- cost of use of machine research output
- amortization of costs in the various uses

(8) *Physical factor*

- form(s) of output
- adaptability to a variety of machines.

Therefore, the above criteria should be carefully considered when evaluating a given system in the light of a need or of a set of given needs. The results of such a study will certainly lead you to conclude whether or not a computer can solve your problems.

An excellent American study:

"Study of mechanization in DOD Libraries and Information Centers - Technical Report AD 640100"

presents the same conclusions in a different form. Another specialist, Mr Frank Allen, has expressed identical opinions.

2. The aim of a documentation center (I am not speaking of a data analysis center) is to provide rapidly, in an easily assimilable form, the documentation indispensable to the user. This assimilable form broadly presents two aspects:

- a rapid description of received documents

- a quick answer to a given question, generally in the form of bibliographic references,

that is to say publication of abstract journals and bibliographic lists. I am leaving aside the reproduction aspect, which does not fall within the scope of my report.

I shall not mention either the advantages of the computer over manual systems: increased easiness in the production of abstract journals, reduced retrieval time for bibliographies, transformation of document loan channels, reduced time, therefore reduced personnel, additional products like, for instance, selective diffusion of information.

To put these advantages in concrete form, an enthusiastic staff, bent on improving itself as well as its working methods, is indispensable; it is therefore necessary to re-train your collaborators. But *what are the categories of documentalists who need re-training?*

3. Good re-training must be adapated to the documentalist's intellectual as well as hierarchic level, and to his specific work.

We shall then consider the following categories:

3.1 "Directors" and senior staff members who advise them as regards the decision to introduce a computer in the service.

3.2 Executives, engineers or scientists qualified in documentation. One can make a distinction between:

- those to be assigned to the computer (analysts, programmers, etc.)

- those who, without being assigned to the computer, are in direct relationship with it (analysts, report writers, lexicographers, etc.)

- those whose work is indirectly related to the computer, or those likely to have their assignment changed.

3.3 Documentalists, assistant-documentalists, librarians, printers, typists, etc.

- those to be assigned to the computer as: operators, console operators, card-punchers, etc.

- personnel likely to be assigned to the computer (replacements, break-down service, etc.)

4. WHO MUST BE RE-TRAINED, AND WHEN MUST RE-TRAINING TAKE PLACE

4.1 To start with, we have mentioned directors and senior staff members who will *use* the services of the computer. Their knowledge of the computer must be synthetic rather than analytic. Lectures or courses laying particular stress on the possibilities offered by the material must be organized for them. Their attention must be drawn to the direct part which they must play in order to derive the best possible benefit from a computer equipped service.

Such lectures or courses should take place *prior to the very decision of* introducing a computer, so that this decision may be made advisedly; in particular, studies on *cost efficiency* must be conducted in order to determine the various operations, or documentation phases, which can be carried out automatically, and efficiently from the cost viewpoint.

Such preliminary re-training of this category of personnel is *indispensable* to adapt the computer, and especially its software, to the needs of the service, and to determine the order in which documentation must be automatized. In other words, these lectures must give emphasis to the elements which determine the choice of the material and the organization of data processing; therefore, it is necessary to have some knowledge of the material and of the programming systems. This re-training phase must be complemented by application courses, when the corresponding application is started, in order that the adjustment of the system may be controlled efficiently and, possibly its operation improved.

4.2 *Secondly*, we have mentioned documentalist-engineers and executives, in particular those to be assigned to the computer as analysts, programmers, or to be directly responsible for automatic data processing services.

There again, very complete re-training must take place prior to the installation of the computer, in order that such personnel may use it immediately and efficiently (computer hours are heavy on a documentation budget).

Personnel directly responsible for automatic data processing services must be re-trained in order to acquire a precise technical knowledge of the material which they will have to manage, but do not need to reach the level of detailed knowledge which will be that of specialized personnel. Usually, senior administrative personnel are quite qualified to occupy such positions.

In view of their turn of mind, specialists used for document analysis and indexing are quite apt to become analysts. These will be in charge of accurately defining and developing the organization planned, then of re-stating problems in the light of the processing means made available to them, and even, possibly, of defining these means.

As regards programming, which is both a technique and an art, specialists of indexing analysis are also, in view of their training, apt to be rapidly re-trained to be application programmers. In other words, they operate and control the computer in advance (they must know in advance what must be done), or else, they speak to the machine in a particular language, in order that this machine may carry out a given task in the least possible time.

Among this category of personnel, one should also consider analysts and documentalist-engineers who stay in their position and continue their work, and whose activity is mainly devoted to feeding the computer indirectly, according to a certain process, and, especially, to establishing a glossary and keeping it up to date. This glossary will be the basis of the whole system.

The glossary, which is coded, is however established and developed by documentalist-engineers (or specialists from the outside) who must also have some notions about the computer and what can be obtained from it. It is they who define and state with accuracy the questions to be put to the machine, in particular for bibliographic research, selective diffusion of information, etc.

Consequently, such personnel must also be re-trained

Besides, it is always useful to have versatile personnel available; therefore, from the same category of specialists one should select those likely to assist, or even replace analysts and programmers if need be. These will also have to be re-trained, but only after the computer has been installed.

4.3 *Thirdly*, you need operators, console operators, card punchers, etc. who must know the general console handling operations (starting, initial loading, resumption of work, stopping, disconnection of computer) as well as the precise handling operations for peripheral units (mounting of bands and discs, use of printers, readers and punchers, etc.)

This offers assistant documentalists, librarians, printers, etc. a new career opportunity since, starting from the position of console operators, they may become later on programmers and therefore get better salaries, etc.

This also applies to typists who, by becoming operators, increase their salaries from 20% to 25% (in France, at least).

As regards this overall category of personnel, re-training, of course, must take place prior to computer installation.

Nevertheless, it should be pointed out that such personnel must take an ability test prior to being re-trained, and, besides, that only persons under 40 should be considered for such positions.

These tests and the age limit restriction will certainly reduce the number of pre-selected candidates, but, on the other hand, have the disadvantage of compelling you to recruit outside.

Now, it is preferable for any personnel in charge of automatization to be also familiar with documentation problems. Mutual comprehension of everyone's respective tasks leads to better results.

5. Complete re-training should include:

- courses and lectures

- constructive courses
- practical assignments and exercises.

5.1 The Directorial Staff, that is to say the Directorate and senior staff members, must attend lectures or lecture series which last approximately 30 days, distributed of course over several months. The subjects covered are broadly as follows;

- general information on conventional material (punched cards, for instance)
- general information on hardware, software and exploitation
- study cases
- role and responsibility of leading staff members in front of the computer
- relations of the computer service within the company
- the processing unit
- external memories
- exploitation systems
- remote processing, etc.

Based on the results of the criteria mentioned at the beginning of this paper, this personnel will be in a position to give instructions on the use of the computer, particularly in certain fields like:

- documentation management
- abstract journals
- indexes
- bibliographies
- selective diffusion of information
- automatic type setting
- and especially, which orientation to follow in the field of equipment to meet the needs of documentation while keeping in mind the cost efficiency of the various documentation operations, that is to say the economic management of the system once installed.

5.2 As for the second category of personnel, that is to say future analysts and programmers, courses will be more complete and detailed than for the first category. They will include, in particular:

Basic data on computers:

- General logic of computers (coding-memory)
- Main input-output units
- Notions on organizational charts
- Languages and principles of exploitation, whatever be the configuration of the machine used.

Programming:

- Technical training for direct work on the machine
- Practical exercises.

Installation and starting:

- General principles on the development of a data processing service:
- Presentation of material to be used:
  - Memory, channels, inputs and outputs, organization of card-indexes, etc.
  - Presentation and comparison of languages
  - Presentation and comparison of exploitation systems.
- General principles governing the implementation of a computer:
  - General analysis methodology
  - Development of analysis and programming data.

The average time allocated to re-training is:

- For analysts: 3 months of pure analysis
  3 months of programming and language for training proper
  15 days for annual re-training
- For already certificated analysts (that is to say, analysts having attended the specialized courses or lecture series organized, as far as we are concerned, by the "Ministry of Finance" (examination in two parts: control test plus oral exam), 3 months of training, plus 15 days of annual re-training are necessary.

5.3 As far as the third category of personnel is concerned, appropriate courses must be attended before and during the installation of the material. Courses last less time but it is only through practical exercises that adequate personnel can be selected. Additional information courses will complement these lecture series.

5.4 Finally, psychological action should especially be taken before and during the installation of the material. Unavoidably, you will note a certain uneasiness within your service, but there again it is indispensable to act with strictness, however with tact. It is difficult for personnel to understand that the computer, far from taking work away from them, gives them more to do; this new work is different, maybe more automatic, but not deprived of interest. As a matter of fact, computers demand a strict discipline in the laying out and development of work, and do not admit of any errors.

6. As a conclusion, certain evident facts could be summed up:

- Consider and analyze your activity thoroughly before introducing your computer, as the re-training which your personnel will have to go through will mean expenses and loss of work (for which you will certainly make up later on).

- Re-training must take place prior to the introduction of the machine.

- Careful selection of the personnel to be re-trained is a vital necessity for a firm. Certain tests are excellent to orient your choice, and you will be surprised at the results obtained by certain candidates.

- Your attention is especially drawn toward the re-training of the leading staff of your service.

PAPER 10

RE-EDUCATING USERS TO AUTOMATED SYSTEMS

by

Carlo Vernimb

Zentralstelle für Dokumentation und Information,
Kommission der Europaischen Gemeinschaften, Luxemburg

SUMMARY

In traditional documentation systems the user performs a search in alphabetical subject indexès by looking up single-term entries. For computerized documentation systems using coordinate indexing and retrieval, precise query formulations are essential. There are different ways of getting the user to make his query more precise. Users' comments are indispensable for evaluating the efficiency of the system. A feedback formsheet is presented, and an example of feedback given. In preparing an SDI profile, a retrospective search should first be performed, which the user must evaluate.

In view of eliminating the documentalist as an intermediary between the user and the system, it can hardly be expected that the user will, just to handle a few queries per year, learn all the retrieval rules the documentalist must bear in mind. On the contrary, the system would have to be developed in a way which enables the applicant to make direct use of it.

SOMMAIRE

Dans les systèmes de documentation traditionnels, l'utilisateur effectue sa recherche en se basant sur les mots-vedette des index-matières. Dans les systèmes de documentation à indexation coordonnée, basés sur l'emploi d'un ordinateur, il est essentiel que les questions soient formulées de façon précise. Les commentaires des utilisateurs sont indispensables pour juger de l'efficacité du système. L'auteur présente le formulaire utilisé pour consigner les commentaires (feedback), et donne un exemple de ce "feedback". Pour établir un profil d'intérêt en vue de la distribution sélective des informations (DSI), il faut tout d'abord effectuer une recherche rétrospective, sur laquelle l'utilisateur devra donner son avis circonstancié.

Quand on veut éliminer le documentaliste en tant qu'intermédiaire entre l'utilisateur et le système, on ne peut guère s'attendre à ce que l'utilisateur, rien que pour traiter quelques questions par an, entreprenne l'étude des règles de la recherche documentaire en ordinateur. Au contraire, le système devra évoluer de telle façon que le demandeur soit capable de s'en servir directement.

RE-EDUCATING USERS TO AUTOMATED SYSTEMS

Carlo Vernimb

A scientist or technician must expend some effort in order to get hold of such scientific publications as he needs for his work. This effort is divided into two operational steps: firstly he must decide which publications he needs, and secondly he must procure the publications. Procurement is easily effected through a library, and this aspect of the problem will not be considered here in any detail. Instead, our interest will be directed toward the effort expended by the scientist to determine which publications he needs.

In traditional documentation systems a scientist will have to search through alphabetical indexes and library card files grouped according to subject, author, etc., and he will have to scan suitable abstract journals. To begin his search, he will need a heading by which to enter the subject indexes. If he does not immediately find an answer to his problem he will try again, under some other entry. After checking a few of the references found, other entries will, as a rule, come to his mind, and he will start searching anew. Such a manual literature search uses single-term entries and applies what I would call "instant feedback". The scientist may start with only a vague idea of his problem; instant feedback will add precision to his query.

In most of the modern computerized documentation systems with coordinate indexing and retrieval, a documentalist acts as an intermediary between the scientist and the system of which the documentalist has to be considered a part. This fact makes it necessary for the documentalist to have a clear idea of the query, so that knowing the scientist's problem, he can add precision to the query.

In our experience with the Euratom Nuclear Documentation System it is just this accurate and clear delimitation of the query which is the most difficult problem. In our request formsheet for literature searches (Fig. 1) which the applicant is asked to use, we therefore give him the following hint:

"Define the subject with enough precision, so that a documentalist specialized in your field but not necessarily acquainted with your particular subject can thoroughly understand the question, and also recognize what would not be of interest to you."

But this advice has proved insufficient to change the habits of the scientists using our system. Most of them still put their queries in the traditional single-term form. We have therefore had recourse to more stringent measures. These measures find their expression in a somewhat aggressive guideline the user will find on the back of the request formsheet (Fig. 2). There it is said:

TELL US WHAT YOU WANT TO KNOW -
SO WE SHALL TELL YOU WHAT IS KNOWN ABOUT THAT.

But do not fling your query at our heads like a brick!

Do not ask for "radioactive strontium compounds" when you can ask more precisely:

> "Which radioactive strontium compounds will in reactor operation be generated within the fuel elements from the strontium produced in uranium fission by chemical reactions with other components of the fuel elements (e.g. carbon, oxygen in UC or UO_2), and so on. ..... The query is of interest because decontamination measures in case of damage to a fuel element will be dependent on the kind of the strontium compounds escaping."

Thus, your query should much less be similar to a title than to a short abstract."

Therefore please do not name us an isolated concept - radioactive strontium compounds - for which you might yourself feel inclined to look in title lists or subject indexes. Concerning radioactive strontium compounds alone we have got more than 3,000 literature quotations stored in the memory of the Computer. Such a mass would overburden you. Our Computer is in need of a meaningful combination of significant concepts, which we shall elaborate from the query formulated by you; this will be the only way for you to obtain the abstracts you are really looking for.

One thing is certain:

> The more accurately you word your query, the more precisely you will know what you want to know, and the better you will learn what our Computer knows about it.

These rather brusque directions have already had some effect. It has become much less frequent for us to have to ask the user to word his query clearly.

Another experience is that a user finds it easier to reveal details and define his needs more exactly on the telephone than in a letter.

There are only a few remarks I would make on the second point: getting user reactions to, and acceptance of, changes in the form and appearance of system output.

Often, the user will consider the results of a computerized documentation system as a complement to the results of the traditional system to which he is accustomed. He does not have to decide straight off that he will restrict himself to the traditional or to the computerized system. If, after having had recourse to the computerized system the user feels that it offers him results comparable to the best results he could obtain from the traditional system, he will come to rely more and more on the computerized system which demands less effort on his part to get access to the desired information.

The rapidity with which he abandons the traditional for the computerized system will depend on the relevance of the items supplied to him, on the completeness of the results of the search, and on the speed of the service. The form of presentation will have a minor but not a negligible influence. Abstracts, copied or in computer-printed form, are preferred; computer printouts of bibliographical references are acceptable; microforms are not looked on with favor. If computer printouts are supplied, those which distinguish between lower case and upper case letters are preferred.

To enable us to evaluate the efficiency of the system and to get advice on increasing its functional value, users' comments are invaluable. In the experimental stage of the Euratom system, queries are processed on the sole condition that the user supplies feedback on the results of our search on a feedback formsheet as follows (Fig. 3):

I do not require any supplementary references.

I should like to receive further references published after 19..

I missed references concerning: . . . . . .

I do not require any references concerning: . . . . . .

I have modified my question as follows: . . . . . .

In addition, he is asked to give the numbers of the references supplied and to say whether they were:

(a) pertinent but already known;

(b) pertinent and new;

(c) not just pertinent but useful;

(d) outside subject limits and so of no interest.

Furthermore, he is asked to quote bibliographical data of pertinent references not supplied to him.

When re-educating scientists in the use of automated systems is considered, this effort of supplying feedback will also count.

Let me give you an example (Fig. 4): A user wanted information on pneumatic devices for control purposes. He wished to have a general view on the subject.

A telephone call to the user brought to light the fact that he was not very familiar with the subject and that he wished to see how far electronic control could be replaced by pneumatic control.

We agreed that a start should be made with a preliminary search and that the query should be subsequently clarified by feedback.

A simple retrieval query (Fig. 5) was formulated with the keyword PNEUMATICS, and the keywords CONTROL or CONTROL SYSTEMS or REMOTE CONTROL.

One hundred of the references retrieved were checked for relevance and 70 were sent to the user.

In his comments (Fig. 6) the user stated that he wished to modify his query to "fluid amplifiers", and that the two references NS 20 08442 and ST 65 29183 were pertinent and new to him, whereas the remaining 68 references, although not entirely pertinent, were useful (I believe, in clarifying his requirements in his own mind at least).

A new retrieval query (Fig. 7) was then formulated with the keyword AMPLIFIERS and the keywords FLUIDS or FLUID FLOW. This query resulted in several additional pertinent references. Of course, it does not always become apparent that after feedback the retrieval query needs to be formulated again, using a completely new set of keywords.

So far my remarks have mainly referred to retrospective literature searches. The time has now come to say a few words on Selective Dissemination of Information (SDI), that is periodical information on recent publications dealing with a certain subject. SDI is intended to replace the traditional browsing through primary and secondary journals. Of an SDI service it will also be true that the speed with which the traditional method is replaced depends on the quality of the service.

For preparing an SDI profile, I should like to recommend the following procedure which has been adopted by Euratom.

At first, each SDI search is handled like a retrospective search, but, if the retrieved references are too numerous, it is restricted to documents published within the past few years. The user is requested to supply feedback. As long as feedback is not received, no SDI profile is prepared.

When feedback has been received, the query is reformulated, in order to reduce machine output as far as possible to such references as the user found to be relevant, ensuring that no relevant item is lost, of course.

Feedback is indispensable because the user could have in the meantime changed, or only made up his mind. This is illustrated by the example on pneumatic control. But there is another reason for reformulating the query, i.e. the experience the documentalist will have gained in checking the relevance of the documents retrieved in the retrospective search.

I should like to give you an example: We received the query "Control of operation of power reactors by computers". We first formulated (Fig. 8)

(CONTROL + CONTROL SYSTEMS + MONITORING)

* (OPERATION + STARTUP + SHUTDOWN)

* (POWER PLANTS + REACTORS)

* (COMPUTERS + DIGITAL SYSTEMS + ANALOG SYSTEMS)

+ means "or"

* means "and"

The primary retrieval terms CONTROL, OPERATION, POWER PLANTS, and COMPUTERS were supplemented by near-synonym terms which could have been used instead in indexing.

With a system volume of 900,000 documents this query resulted in 180 documents retrieved, 40 of which proved to be relevant; consequently, relevance ratio was calculated as 22%.

When checking the relevance of the documents retrieved, the documentalist detected (1) that the terms DIGITAL SYSTEMS and ANALOG SYSTEMS contributed no relevant items and (2) that many references were retrieved in which the computer was used for calculations (indicated by the keyword NUMERICALS), or for simulating the behavior of reactors. Furthermore, he realized that the indexers might have used only one term of the first two lines of the query formulation to index a relevant document. "Control of power reactors by computers" could as well be a title for a relevant document as "Operation of reactors by computers". As a consequence, the first two lines were combined, which results in the new query formulation (Fig. 9):

(CONTROL + CONTROL SYSTEMS + MONITORING + OPERATION + STARTUP + SHUTDOWN)

* (POWER PLANTS + REACTORS)

* COMPUTERS

*- (NUMERICALS + SIMULATORS)

*- means "and not"

This query yielded 660 documents, 480 of which proved to be relevant - a relevance ratio of 73%. Thus, by reformulating the query, the number of pertinent items retrieved increased by a factor of 12.

The average relevance ratio for retrospective searches turns out to be more than 40% in the Euratom System. Making use of the documentalist's experience in screening the results of the retrospective search, and taking the user's feedback into account, it was possible to obtain an average relevance ratio of about 80% for SDI searches. By the way, I do not know of any large documentation system in operation in which such a high relevance ratio is being achieved.

The user is expected to accept a certain amount of "noise". Say, 20% of irrelevant documents might be acceptable. The percentage could even be higher for small shipments, because the user's workload for checking relevancy would be small.

We have got one-and-a-half years' experience with now about 300 SDI profiles applied to monthly batches of 10,000 documents, and resulting in an average of 10 abstract copies to be sent to each user without any further checking of relevancy. With four or five exceptions, all users have accepted this standard of service. So far there has been virtually no change of profiles. But several users cancelled their profiles and desired new ones to be established to meet changes in their fields of activity.

In this diagram (Fig. 10) the user's effort in applying to a traditional documentation system, with its alphabetical subject indexes, library card files and abstract journals, is compared with his effort in availing himself of a computerized documentation system with coordinate indexing and retrieval. In the first case, the effort consists in searching through alphabetical indexes via single-term entries (headings), with application of instant feedback to make the query more precise. In the second case, the effort consists in formulating a precise query, commenting on results, and rephrasing the query if necessary; feedback is thus considerably delayed as compared with the traditional method. For the computerized documentation system, the documentalist must complement the user's effort by interpretation and formulation of the query, by logical coordination of search terms, by checking the results, and by reformulation of the query if necessary.

The fact that the user's effort will decrease as he changes over from the traditional to the computerized system will at first please him, but his enthusiasm will be somewhat cooled by the need to adjust to a new technique. Even the higher yield of useful documents obtained by the computerized system will not deceive him as to the fact that he must rely upon the system documentalist as an intermediary, and that he cannot start a search possessing only a vague idea of the problem. You may think that obliging the user to express his needs more clearly will in any case be to his profit, that he will understand his own problem better. But if this is not possible (because he is not clear in his own mind about his problem) the single-term-entry method with instant feedback would offer some advantage. And who can say whether the latter method is not better suited to the human way of thinking?

The problem is how to do away with the system documentalist. This could be done by shifting the borderline between the user's and the documentalist's efforts to the right, or by shifting the borderline between the documentalist and the system to the left, or even by shifting both borderlines towards each other.

We had some disappointing experiences in getting the user interested in query formulation. For such a simple query as "Isotope enrichment for uranium 235 by ultracentrifuges" (Fig. 11), one user looked up the Euratom Thesaurus and proposed the formulation:

ISOTOPES * ENRICHMENT * URANIUM 235 * ULTRACENTRIFUGES

But an indexer assigning URANIUM 235 should never add the term ISOTOPES. Moreover, according to a specific indexing rule, the keyword ISOTOPE SEPARATION would have to be used. The correct formulation should have been:

ISOTOPE SEPARATION * (URANIUM 235 + URANIUM ISOTOPES) * ULTRACENTRIFUGES

The user's formulation would have yielded virtually no answer - and this is a simple example only. It can hardly be expected that a user will, just to process a few queries a year, learn all the indexing and retrieval rules the documentalist has to bear in mind.

To try and make the system documentalist superfluous, we must, in my opinion, shift the right borderline considerably to the left, and the left borderline only a little to the right. Shifting the right borderline to the left means, first of all, direct access to the computer. Direct access is being used in several experimental documentation systems. But, as far as I can see, and as far as performance is concerned, it will not be possible to replace the documentalist by direct access alone.

We can count on the ability of the user to judge between pertinent and irrelevant documents. What we need is a system which will present to the user - by way of direct access - documents or abstracts of documents on which he can base his judgment concerning relevance, which is then directly fed back into the system.

Let me sketch roughly the interaction between user and system, without the documentalist as intermediary:

1. The user feeds his query in natural language, via a console, into the computer.

2. The query is indexed automatically, i.e. keywords are derived from it.

3. The system presents these keywords, as well as semantically related keywords, via a television screen, to the user who chooses the most appropriate and coordinates them by the logical operation "and" and "or".

4. The system then indicates the number of answers that would be retrieved by the proposed formulation, and the user is given the possibility to reduce the number of answers by adding a keyword (with "and") representing an additional condition, or to increase the number of answers by leaving away less significant keywords.

5. Bibliographical data or abstracts of the document retrieved are shown up on the screen.

6. The user checks these documents for relevance until he has obtained, say 4 clearly pertinent and 4 characteristically irrelevant documents.

7. The reference numbers of the clear "hits" and of the characteristic "noise" are fed back into the system.

8. In such a relevance feedback procedure the remaining retrieved answers on the query are ranked in the order of their pertinence probability. A document will have high pertinence probability when its indexing is similar to that of documents already found to be pertinent.

9. The user checks the ranked documents for relevance, on the screen, if the number is limited, on a computer printout supplied to him, if the number is higher

This procedure can run through a number of loops. In addition, the query can be automatically loosened by adding semantically similar terms as alternatives to the original query time. I should like to remind you of the terms STARTUP and SHUTDOWN, used as alternatives to OPERATION in our example. The query can also be "tightened" automatically by formulating it out of the keywords initially used to index the pertinent documents.

By tightening and loosening the query formulation, always adjusting it by checking for relevance, and by using the relevance feedback procedure, it should be possible to focus the query formulation on the pertinent documents contained in the system. This is a complex procedure, but essential parts of it have already been tested, and there is reasonable hope that one day not only the documentalist will be made superfluous as an intermediary between the user and the system but also it will be made possible for the user to start his search with only a vague idea of his problem and to add precision to his query by instant feedback

COMMISSION OF THE EUROPEAN COMMUNITIES
Directorate General "Dissemination of Information"
CID
Documentation

RS-

Date:

LITERATURE SEARCH

REQUEST

1. **FIRM OR ORGANIZATION:**
Requester's name
Address for correspondence:
Telephone/telex No.: Extension No.:

2. **SUBJECT OF THE SEARCH:**
(Define the subject with enough precision so that a documentalist specialized in your field but not necessarily acquainted with your particular subject can thoroughly understand the question, and also recognize what would not be of interest to you). To help us pinpoint your problem, please mention a few relevant documents and, if possible, supply a copy of their first page.

3. **PURPOSE OF THE SEARCH:**

a) I would like to have a **general view** on the subject of my request (to become acquainted with a new field or an urgent technical problem). ☐

b) I would like to **complete my own bibliography** in my specialized field or check its completeness. ☐

I already know ____________ relevant publications.

The search should be limited to documents published after ____________
For this period, I expect to receive ____________ references.

c) Can you quote me any documents published prior to ____________ and possibly impairing the novelty of the inventive idea exposed under Item 2? ☐

4. **ANY OTHER REMARKS:**

5. Please send me one copy of every abstract pertinent to my query. I am aware of the fact that I may use such copies only for personal ends and not divulge them (handing over of copies to staff members working with my own employer is not deemed to be divulgation), and that the author is entitled to a fair remuneration if the reproduction serves commercial purposes.

____
(signature)

Figure 1

Tell us what you want to know – so we shall tell you what is known about that

But do not fling your query at our heads like a brick!

Dot not ask for «*radioactive strontium compounds*»,

when you can ask more precisely:

> «***Which radioactive strontium compounds will in reactor operation be generated within the fuel elements from the strontium produced in uranium fission by chemical reactions with other components of the fuel elements (e.g. carbon, oxygen in UC or*** UO_2***), and by which reaction mechanisms will such strontium compounds originate? It is said that strontium oxide has so far been traced. A pertinent publication, however, is not known to us. The query is of interest because decontamination measures in case of damage to a fuel element will be dependent on the kind of the strontium compounds escaping***».

Thus, your query should much less be similar to a title than to a short abstract.

Therefore please do not name us an isolated concept – ***radioactive strontium compounds*** – for which you might yourself feel inclined to look in title lists or subject indexes. Concerning radioactive strontium compounds alone we have got more than 3000 literature quotations stored in the memory of the Computer. Such a mass would overburden you. Our Computer is in need of a meaningful combination of significant concepts, which we shall elaborate from the query formulated by you; this will be the only way for you to obtain the abstracts you are really looking for.

One thing is certain:

> The more accurately you word your query, the more precisely you will know what you want to know, and the better you will learn what our Computer knows about it.

We shall not tell anybody what you want to know – but we shall tell everybody what you have published

At the present time, searches are carried out free of charge, on condition, that is, that you give us your comments on the information received from us («feedback»). – We do not accept any responsibility concerning the accuracy, completeness or utility of the documents mentioned by us.

Address: CID/ENDS, 29, rue Aldringer, Luxembourg – Tel.: 29241 – Telegraphic address: «COMEUR Luxembourg» – Telex: «COMEURLUX 423 or 446»

Figure 2

RS –

COMMENTS

ON LITERATURE SEARCH

These comments are the only compensation we ask from you for processing your query.

Subject of search:

1. I do not require any supplementary references ☐
2. Please keep me regularly informed on this subject ☐
3. I should like to receive further references published after 19 ....
4. I missed references concerning:

5. I do not require any references concerning:

6. I have modified my query as follows:

Please indicate, in each column, up to ten complete reference numbers relating to clear "hits" (Columns "a" and "b"), and to clear "noise" (Column "d"). In addition to the reference numbers give the total number of references per column.

Among the references supplied, the following were:

a) pertinent, but already known	b) pertinent and new	c) not just pertinent, but useful	d) outside subject limits, and so of no interest	e) references not supplied, although pertinent (Quote bibliographical data in full. Use a separate sheet if necessary)

Date : Signature :

Address: CID/ENDS, 29, rue Aldringer, Luxembourg – Tel.: 292 41 – Telegraphic address: «COMEUR Luxembourg» – Telex: «COMEURLUX 423 or 446»

Figure 3

COMMISSION OF THE EUROPEAN COMMUNITIES
Directorate General "Dissemination of Information"
C I D
Documentation

RS-

Date:

LITERATURE SEARCH

REQUEST

1. **FIRM OR ORGANIZATION:**
   Requester's name:
   Address for correspondence:
   Telephone/telex No.: Extension No.:

2. **SUBJECT OF THE SEARCH:**
   (Define the subject with enough precision so that a documentalist specialized in your field but not necessarily acquainted with your particular subject can thoroughly understand the question, and also recognize what would not be of interest to you). To help us pinpoint your problem, please mention a few relevant documents and, if possible, supply a copy of their first page.

   Pneumatic devices for control purposes

3. **PURPOSE OF THE SEARCH:**

   a) I would like to have a **general view** on the subject of my request (to become acquainted with a new field or an urgent technical problem). [X]

   b) I would like to **complete my own bibliography** in my specialized field or check its completeness. [ ]

   I already know ______________ relevant publications.

   The search should be limited to documents published after ______________.
   For this period, I expect to receive ______________ references.

   c) Can you quote me any documents published prior to ______________ and possibly impairing the novelty of the inventive idea exposed under Item 2? [ ]

4. **ANY OTHER REMARKS:**

5. Please send me one copy of every abstract pertinent to my query. I am aware of the fact that I may use such copies only for personal ends and not divulge them (handing over of copies to staff members working with my own employer is not deemed to be divulgation), and that the author is entitled to a fair remuneration if the reproduction serves commercial purposes.

____
(signature)

Figure 4

Pneumatic devices for control purposes

PNEUMATICS * (CONTROL + CONTROL SYSTEMS + REMOTE CONTROL)

Figure 5

AMPLIFIERS * (FLUIDS + FLUID FLOW)

Figure 7

Control of the operation of power reactors
by computers

(CONTROL + CONTROL SYSTEMS + MONITORING)

* (OPERATION + STARTUP + SHUTDOWN)

* (POWER PLANTS + REACTORS)

* (COMPUTERS + DIGITAL SYSTEMS + ANALOG SYSTEMS)

+ means "or"	180 answers
* means "and"	40 hits
	22 % relevance ratio

Figure 8

(CONTROL + CONTROL SYSTEMS + MONITORING + OPERATION
+ STARTUP + SHUTDOWN)

* (POWER PLANTS + REACTORS)

* COMPUTERS

*- (NUMERICALS + SIMULATORS)

*- means "and not"	660 answers
	480 hits
	73 % relevance ratio

Figure 9

RS –

COMMENTS

ON LITERATURE SEARCH

These comments are the only compensation we ask from you for processing your query.

Subject of search:

1. I do not require any supplementary references
2. Please keep me regularly informed on this subject
3. I should like to receive further references published after 19 ....
4. I missed references concerning:
5. I do not require any references concerning:
6. I have modified my query as follows:

Fluid amplifiers

Please indicate, in each column, up to ten complete reference numbers relating to clear "hits" (Columns "a" and "b"), and to clear "noise" (Column "d"). In addition to the reference numbers give the total number of references per column.

Among the references supplied, the following were:

Quote bibliographical data in full. Use a separate sheet if necessary

a) pertinent, but already known	b) pertinent and new	c) not just pertinent, but useful	d) outside subject limits, and so of no interest	e) references not supplied, although pertinent
	NS 20 08442	the remaining		
	ST 65 29183	68 references		

Date : Signature :

Address: CID/ENDS, 29, rue Aldringer, Luxembourg – Tel.: 292 41 – Telegraphic address: «COMEUR Luxembourg» – Telex: «COMEURLUX 423 or 446»

Figure 6

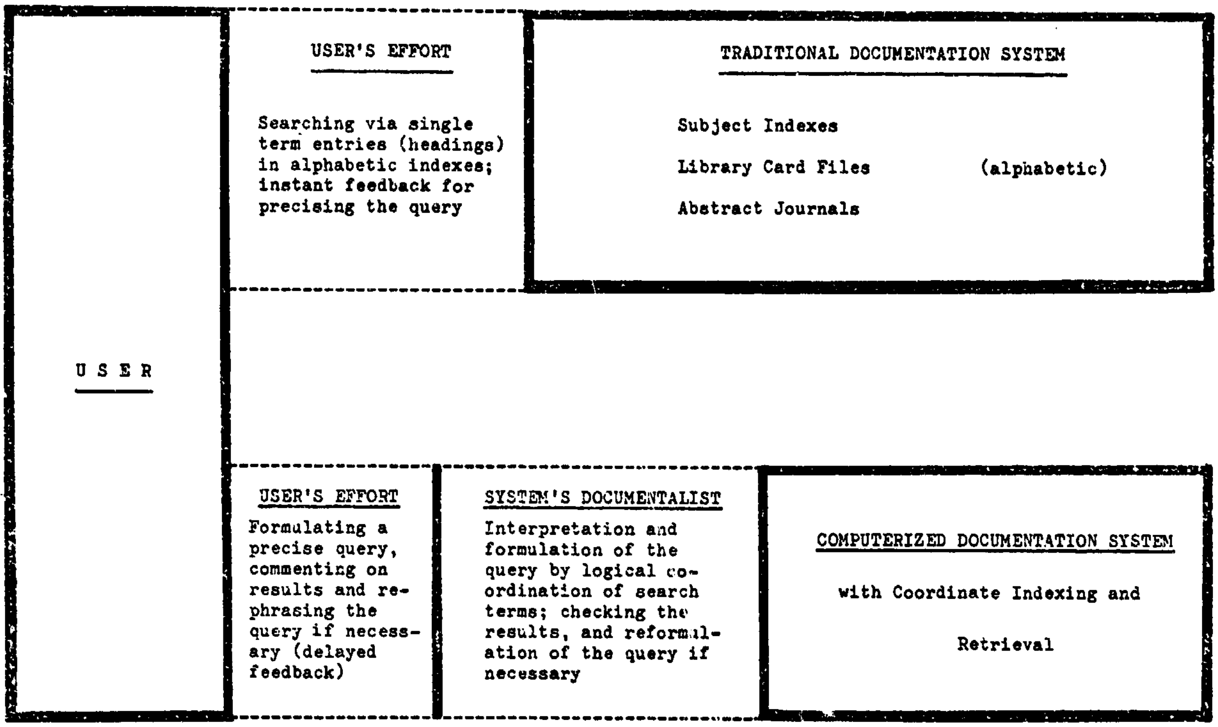


Fig. 10 User's effort in applying to documentation systems

Isotope enrichment for uranium 235 by ultracentrifuges

ISOTOPES * ENRICHMENT * URANIUM 235 * ULTRACENTRIFUGES

ISOTOPE SEPARATION * (URANIUM 235 + URANIUM ISOTOPES) * ULTRACENTRIFUGES

Figure 11

PAPER 11

EVALUATION AND COST ANALYSIS OF COMPUTERIZED INFORMATION ANALYSIS CENTER OPERATIONS

by

J.R. Buchanan

Nuclear Safety Information Center,
Oak Ridge National Laboratory,
Oak Ridge, Tennessee, USA

SUMMARY

Computerization can enable an intermediate-size information analysis center (2,000 to 12,000 accessions per year) to most efficiently use its professional and clerical staff and to prepare for growth. In order to make effective use of its funds and prepare for expansion or increased scope, it is important to plan ahead for the collection of operating data that will permit cost analysis of the various production operations connected with collecting, processing, and outputting the services offered. Measurement of user response and effectiveness of the services are vital to this analysis. The experience of the Nuclear Safety Information Center and other information analysis centers will be used to illustrate how this has been done.

SOMMAIRE

Le recours à l'ordinateur peut permettre à un centre d'analyse d'informations d'importance moyenne (de 2,000 à 12,000 admissions par an) d'obtenir le meilleur rendement possible de son personnel technique et administratif, et de préparer son expansion. Il est important, si l'on veut utiliser efficacement ses fonds et se préparer à s'accroître ou à étendre le champ de ses activités, de prévoir à l'avance le recueil d'informations sur le fonctionnement, qui permettront de procéder à l'analyse du coût des diverses opérations de production liées au recueil, au traitement et à la production des services offerts. La mesure de la réaction de l'utilisateur et de l'efficacité des services est, pour cette analyse, d'une importance vitale. Pour illustrer comment ceci a été réalisé, l'auteur rendra compte de l'expérience du Centre d'Information de la Sécurité Nucléaire ainsi que d'autres centres d'analyse d'informations.

EVALUATION AND COST ANALYSIS OF COMPUTERIZED INFORMATION ANALYSIS CENTER OPERATIONS*

J. R. Buchanan

An information analysis center may be thought of as a manufacturing concern where information and/or data is processed in various forms during well-defined production operations. Finished products are distributed to users who originate requests for relevant material on particular subjects. The management function of a center is therefore somewhat analogous to that of a manufacturing firm, i.e., the enterprise is concerned with the same basic functions of purchasing, production, service, finance, administration, and distribution[1]. It is in this fashion that evaluation of processes and cost analysis becomes important. In this case, we are particularly interested in their application to computerized activities.

The significance of computerized operation is that it can enable management to make most effective use of its personnel and the single most important facet of the operation of an information analysis center (IAC) is its staff of experienced scientists and engineers. At the Nuclear Safety Information Center (NSIC)[2,3] and other similar organizations, these technical specialists' duties include preparing review articles and reports, answering technical inquiries, and cataloging reference material.

While the principal functions of the first two tasks are not computerized, they involve the greatest portions of the time of the technical staff and are vital to the operation of an effective IAC. Computerization enables such a center to most efficiently use its professional and clerical staff and prepare for growth. Benefits that can result from mechanization include:

Increased service

Faster service

More search flexibility

Greater scope expansion flexibility

Lower personnel costs

In this paper the experience of a number of organizations including NSIC will be utilized to illustrate how cost data may be collected and how effectiveness or user satisfaction may be measured. Some comparative statements on manual and computerized operations are presented followed by a summary of cost analysis criteria and evaluation criteria. Next comes a discussion of cost and effectiveness practices followed by a summary of real problems of computerized operations that have been identified by IAC managers.

Comparison of Manual and Computerized Operations

Since the subject of this paper is the cost and analysis of computerized operations, it is assumed that this audience either is involved in such an activity or is strongly considering computerization. For those still using a manual system or only recently converted to the computer, it is of some interest to compare the two modes of operation. Unfortunately, facilities with computerized operations that formerly were manual have generally not attempted to compare the two methods except in superficial ways. The Booz, Allen study of Mechanization in DOD Libraries and Information Centers[4] did contain some interesting qualitative evaluations of the effectiveness of their own mechanized systems by several facilities' staffs. (Although some of the data was from library operation, it is included here because of its general interest and the scarcity of this type of information.)

1. At the Picatinny Arsenal, a computerized system took an average of 0.15 hour per item for keypunching, sorting, pulling from shelves, preparing for mail, discharging refiling returns, and processing on the computer. The manual system required 0.18 hour per item.

2. The Technical Information Reference Branch at Wright-Patterson AFB has been able to produce by computer a regular accessions list twice per month with no increase in effort or manpower. In the eight years preceding mechanization, the Branch was able to produce manually only one list every other month.

3. The US Naval Postgraduate School Library now devotes about 10 to 15 minutes each to 300 student requests per month with a computer system, saving about [illegible] man hours per month. Formerly 1½ hours was required for professional assistance with an average manual search.

---

* Research sponsored by the US Atomic Energy Commission under contract with the Union Carbide Corporation.

4. The Air Force Materials Laboratory Library Staff found that reference searches produced by the computer caused a reduction of irrelevant information of over 56% and incurred less than a 5% loss of relevant information.

5. The staff at the Redstone Scientific Information Center was able to perform a greater volume of library service at a saving in staff time through computerization. In 1962 a staff of 67 was required to serve 4000 users. After mechanization in 1965, there were 8300 users, while the staff had been decreased to 43 members, with no increase in overall workload on the staff.

6. The Infrared Information Analysis Center was able to answer five times as many requests after computerization as before. Prior to going on full-time computer use, IRIA personnel required two to three days to prepare a bibliography for a visitor. The time now required averages one and one-half hours.

7. In 1964 the Foreign Technology Division Library at Wright-Patterson conducted a controlled experiment on a data base of 100 documents. The computer search based on 30 user profiles retrieved an average of 85 documents. The manual searches yielded 45 documents. Of those 45, about 15%, or seven documents, had been missed by the computer search. Relevance levels were 58% on the computer search and 63% on the manual search.

8. Conversely, all four of the information centers at the Battelle Memorial Institute that were included in the survey consider that their manually produced and searched files of extracts are preferable to a computerized system. However, computerization of the center's printed outputs will be considered for time-saving features where it is felt to be economically advantageous.

Cost Analysis Criteria

The most important measures of system analysis are (1) cost, (2) performance, and (3) benefits and their interrelationships. Cost describes the expense of the system in dollars. System performance is determined by such attributes as accuracy, speed, quality, extent of coverage, etc. Benefits describe the consequences of the system in terms of (1) its value, determined from such factors as number of references used per year or duplication of research avoided; (2) how the system affects the behavior of persons; or (3) how it affects related systems. A theoretical benefit-cost-performance described by Murdock and Liston[5] is found by plotting a benefit to cost ratio against percentage of maximum possible performance. This, of course, requires all benefits of the system to be convertible to dollars and this is not always possible.

While libraries do not generally maintain cost data, information center management conversely is generally very concerned about costs and maintains detailed records of current operating costs. However, even in these cases, historical development costs are usually not available. Booz, Allen[4] felt that incomplete cost records and insufficient concern relative to cost versus usefulness and value was probably the most detrimental influence on the development of many of the systems they studied.

The following criteria relative to cost analysis seem basic:

1. *Purpose of Collecting Cost Information* - The purpose of assembling cost information is to enable management to compare fund consumption with allocation. The allocation is usually specified by a budget which should reflect the judgment of management concerning the relative value of each activity, both tangible and intangible. For computerized activities, this means including the value of such elements as increased timeliness, flexibility, and quality; broader and/or deeper search capability; greater file detail; increased user satisfaction, etc. These same elements should also be used in measuring the operating systems effectiveness thereby providing a two-way check - budget versus cost and expected value versus actual effectiveness.

2. *Budget Revision* - The budget will be subject to revision based upon the availability of funds and changing views on the value of system elements and output. The budget should realistically reflect the anticipated value of the developing system as well as the available funds. Costs should be recorded as they accumulate, including reasonable estimates for those services that are supplied "free". They can then be related to the budget, the costs of the old system, and the effectiveness of the new system. The budget may then be modified to direct the available funds to those areas that will yield the greatest return in terms of the goals established.

3. *Cost Elements* - The elements that make up the overall cost of computerized processes include:

   Staff salaries
   Materials
   Equipment purchase or rental
   Equipment space requirements
   Maintenance
   Computer time
   Programming
   Contractor services
   Development of new skills

These cost elements after being assembled must be related to various units of the operation, cost per acquisition cataloged, per search, per dissemination, etc. These figures then permit a cost analysis of the computerized processes in terms of budget and, combined with the operation statistics, a determination of cost effectiveness. Finally, these figures can be maintained from year to year to develop an historical cost basis for comparisons, decisions regarding future budgets, staff assignments, and changes in services offered to the users.

King[6] has identified several papers on computer and programming costs of interest to system designers and evaluators - including Erikson[7], Marron and Synderman[8], Nelson[9,10], Tupac[11] and Weinwurm[12].

Evaluation Criteria

In introducing the subject of evaluation, Salton[13] has shown that it is necessary to make a number of distinctions because (1) there are different kinds of systems (operational, experimental, and laboratory), (2) there are different viewpoints concerning the system (user, manager, operator), and (3) systems are composed of many subsystems. Therefore, many different variables are identified which affect system evaluation. Salton identifies the following factors as among the most important:

1. *User Population* - Type of user, rate of requests, etc.

2. *Collection* - Coverage of collection, type of document available at input, reliability of abstracts, etc.

3. *Indexing* - Type of indexers, level and accuracy of indexing, depth of indexing required, and complexity of indexing language required.

4. *Analysis and Search* - Type of searching, power and complexity of search mechanism, search effort required, accuracy of search, etc.

5. *Equipment and Input-Output* - Type of store, type of in-out equipment, and type and form of output.

6. *Operating Efficiency* - Cost considerations service problems, time lag, and response time.

Treu[14] in a literature survey found six different test and evaluation methods that had been applied to both individual systems and comparison of systems. These are:

1. Search and system response procedure involving a request, interpretation of same, query formation, and evaluation of results.

2. Out-of-system-service procedure for testing one or more selected system elements or functions.

3. The questionnaire and interview technique for determining user requirements and evaluating system service.

4. Observation of the operating system for such things as workload, staff performance, and system operation.

5. Simulation by such means as mathematical models.

6. Multiple approach using two or more of the above methods.

Many feel that the overriding criterion for evaluating effectiveness is the ability of a system to satisfy the user and that the principal aims of the other participants, such as managers and operators, are the same. Cleverdon[15] recognizes six main user criteria which affect his satisfaction:

1. The *coverage* of the collection or the extent to which the system includes relevant matter.

2. The *time lag* or average interval between the time the search request is made and the time an answer is given.

3. The form of *presentation* of the output.

4. The *effort* involved on the part of the user in obtaining answers to his search requests.

5. The *recall* of the system is the proportion of relevant material actually retrieved in answer to a search request.

6. The *precision* of the system is the proportion of retrieved material that is actually relevant.

Collection of Cost and Effectiveness Information

Most of the papers published on the evaluation of information systems dwell on measures of the recall and precision of the information storage system. While these are important tests, it is possible that they have been over-emphasized relative to other means of test and evaluation because they are usually easier to measure and report in a meaningful fashion. In this section we look at other methods such as feedback and cost measurement.

User Response

MSIC has employed various means of getting feedback from the user and measuring user satisfactions. Each set of SDI abstract cards mailed from the Center has a feedback card that has been prepared by the computer. A typical card is shown in Figure 1. The system has 1700 users and output is prepared and sent out bi-weekly.

Profile adjustments are made on the basis of feedback received from the users. The procedure has been previously described[16].

Questionnaires have also been used by NSIC to evaluate services. A typical one is shown in Figure 2. Dugger and Klinger[17] have attempted to determine by questionnaire the value of information received from an information system in terms of dollars or man hours saved by preventing duplication of effort, loss of time, etc. While users of information systems frequently save time and money in this fashion, it is almost impossible to obtain meaningful numbers on the exact amount of the saving.

Cost Evaluation

The cost and budgeting experience of NSIC and the Aerospace Research Applications Center (ARAC) at the University of Indiana will now be discussed.

NSIC Budget - All projects at Oak Ridge National Laboratory (ORNL) have their operating costs tabulated each month by computer. A copy of the Operating Cost Report for NSIC for May 1969 is shown in Figure 3. While there is a detailed breakdown of personnel cost, the computer expenses are reported as a single line item as "EDP(CDP & ORNL)". More detailed charges for data processing, programming, and remote consoles are reported in a separate form.

Using cost figures developed in this fashion, it has been possible to determine the expense of doing basic operations and to budget the various activities. At NSIC the budget for FY-69 was allocated as shown in Table I. All computer charges are lumped under one item so that the other entries such as SDI are for expenses at NSIC exclusive of computer costs. The expense of doing various computerized operations on the IBM-7090 is shown in Table II. Comparable figures for the IBM-360 are not available due to limited operating experience to date in the new production mode[17]

NSIC Input - Abstracting and indexing costs associated with technical editing and clerical efforts was \$11.10 per abstract in FY-69. It cost an additional \$1.40 to store each abstract on the computer for a total cost of \$12.50. We have no figures from comparable systems for direct comparison; however, one large abstracting service unofficially reported that it cost them \$20 per abstract to index, abstract, and produce finished copy for an abstract journal.

NSIC Indexed Bibliography - NSIC reports are sold for \$3 per copy by the Clearinghouse for Federal Scientific and Technical Information (CFSTI). Since the Indexed Bibliographies are issued bi-monthly, this made a yearly set cost \$18 for a total of 4820 abstracts in FY-69. The production expense of \$4.58 per copy is essentially all associated with printing since the \$50 computer cost amounts to only 8 cents per copy for the present print run of 628. The printing costs were paid from ORNL overhead funds in FY-69 but will be charged directly to NSIC in FY-70. None of the abstracts input costs have been charged against the Bibliography since it is thought of as a by-product primarily for library use. (The abstracts must be in the storage file anyway so that the SDI and retrospective search programs can function.) For comparative purposes we note that a commercial publication *Chemical Abstracts* had an annual subscription rate of \$1550 in 1968 when it contained 232,704 abstracts for an average cost to the subscriber of 0.67 cents per abstract. (If the cost were set for full recovery of printing expenses, the NSIC charge would be 0.57 cents per abstract.)

NSIC-SDI - At present there is no charge for the SDI service although NSIC has surveyed its users to help determine what it is worth to them. One hundred questionnaires were mailed at random asking what they would be willing to pay for the service. A total of 69 forms were returned with the following results.

1. 42 willing to pay
   20 not willing to pay
   7 returned form but did not answer question
2. Fees checked were — and write-ins were

Fees checked were	and write-ins were
23 - \$ 25/yr	1 - \$ 5/yr
9 - \$ 50/yr	2 - \$ 10/yr
3 - \$100/yr	1 - \$400/yr

The cost to NSIC of selecting, sorting, printing, and mailing SDI is about 7 cents per abstract and the average user receives about 15 abstracts every one-to-two weeks. The yearly cost is \$51.05 including the expense of profile maintenance and billing and fee collection. Input costs add an additional yearly charge of \$4.32.

The USAEC has yet to reach a decision as to its policy on charges for information center services. An AEC Technical Information Panel Task Force has recommended that these services not be considered in the general category of "Goods and Services" and that the sponsoring program division recognize the need to budget all information activity costs except those of an unusual nature for unusual needs.

NSIC Consultation and Retrospective Searches - During the past year, NSIC responded to 609 technical inquiries. The average non-computer related cost of replying to each was \$75.70. Most of the expense is a direct result of the amount of technical consulting time required. About 2/3 of the inquiries also required backup references. This resulted in over 400 retrospective searches. The computer associated cost of a search, as we have seen earlier, . \$25 when they are processed in batches.

ARAC Experience - At ARAC, a computerized cost system was developed by Helmkamp[1] and operated for a 3-month period to show that it is possible to design a sound cost accounting system to meet the specific requirements of an information center. The components of the system were designed specifically for ARAC. These included such items as labor time records, work orders, labor performance reports, and service cost reports. An abbreviated monthly cost performance report is shown in Table III. Using the various basic cost data, cost distributions were developed for retrospective searches and SDI as shown in Table IV.

There has been no attempt to compare the cost figures collected by NSIC and ARAC since a good deal of normalization would be required to put them on the same basis. The point to be made is that data were collected and reported in the fashion most useful to each particular management group. It is obvious from the experience of these two centers that meaningful cost and effectiveness data can be collected. The essential ingredients are to plan early for its collection and to design a system that allows for easy interpretation of results.

Problems Related to Computer Operations

Some general problems related to use of computers by IACs have been identified since they ultimately have an effect on costs and effectiveness. Typical problems are as follows:

1. *Computer Service* - Some groups are forced to rely on other organizations for computer services. This can result in service that is erratic and sometimes unavailable for extended periods of time.
2. *Changeover* - Bringing a computer storage and retrieval system into operational use is greatly complicated by the problems of file conversion. Computer searching is effective only if a substantial part of the file to be searched is accessible to the computer. However, the creation of a complete machine file often initially exceeds the funds available and the file conversion process may have to be stretched out over a long period. This experience can be frustrating and discouraging to both the user and the searcher[4].
3. *Staffing* - Information facilities frequently consider themselves to be understaffed for the job expected of them. Some of the difficulty in bringing a computer process into operation can be attributed to insufficient staff for organizing the system and preparing material for computer input[4].
4. *Organization* - Centers reporting to research and development staff management generally feel that they have much better support for their computerization efforts (as well as their other activities) than do those reporting to administrative management[4].
5. *Expense of Computerized Activities* - The amount of money spent on computerized activities by NSIC has increased each year. This presents a problem because of increasingly tight budgeting times. It is, however, a tribute to the success of the computer programs in that more use has been made of them than was ever envisioned when the work was started. Also, to perform as many activities manually as are currently being done on the computer would certainly cost considerably more. At present, about one-third of NSIC's budget is consumed in computer program development and operation.
6. *Desire to do More on Computer* - The desire to do even more on the computer frequently exceeds the programming manpower and time available to provide needed programs or revisions. It becomes necessary to establish priorities for even the small jobs. Again, at NSIC this is considered a tribute to the success of the existing programs.

Conclusion

The most efficient use of the professional and clerical staff of an IAC can be made by a balanced program of mechanization. However, it is very important to plan early for the collection of operating data that will permit cost analysis of the various center operations. This will enable the managers of the operation to make most effective use of their funds and plan for the future.

REFERENCES

1. Helmkamp, John G. — *Managerial Cost Accounting for a Technical Information Center.* American Documentation, April 1969, pp. 111-118.
2. Buchanan, J.R.; Cottrell, Wm.B. — *Operating Experience of the Nuclear Safety Information Center March 1963 - March 1965.* USAEC Report ORNL-TM-1136, Oak Ridge National Laboratory, May 17, 1965.
3. Buchanan, J.R.; Cottrell, Wm.B. — *A Summary of NSIC Activities 1963-1967.* USAEC Report ORNL-NSIC-46, Oak Ridge National Laboratory, September 1968.
4. - — *Study of Mechanization in DOD Libraries and Information Centers.* AD-640 100, Booz, Allen Applied Research Inc., September 1966.
5. Murdock, John W.; Liston, David M., Jr — *A General Model of Information Transfer: Theme Paper 1968 Annual (ADI) Convention.* American Documentation, Vol. 18, Oct. 1967, pp. 197-208.

6. King, Donald W. — *Design and Evaluation of Information Systems.* Chapter 3 of Annual Review of Information Science and Technology, Vol.3, 1968.

7. Erikson, Warren, J. — *An Analytical Cost Comparison of Computer Operating Systems.* System Development Corp., Santa Monica, Calif., TM-3525, June 30, 1967, p.214.

8. Marron, H.; Snyderman, M. — *Cost Distribution and Analysis in Computer Storage and Retrieval.* American Documentation, Vol.18, July 1967, pp.162-164.

9. Nelson, E.A. — *Management Handbook for the Estimation of Computer Programming Costs.* System Development Corp., Santa Monica, Calif., TM-3225/000/01, March 20, 1967.

10. Nelson, E.A. — *The VOPIR Technique for the Evaluation of Computer Augmented Administrative Systems.* System Development Corp., Santa Monica, Calif., May 17, 1967, p.71.

11. Tupac, J.D. — *An Approach to Software Evaluation.* RAND Corp., Santa Monica, Calif., Report No. P3581, AD-651 812, April 1967, p.13. Presented to the American Management Association's Annual EDP Conference, New York, March 6, 1967.

12. Weinwurm, G.F. — *On the Economic Analysis of Computer Programming.* System Development Corp., Santa Monica, Calif., SP-2713, August 1967, p.37.

13. Salton, Gerard — *Automatic Information Organization and Retrieval.* McGraw-Hill Book Company, 1968, p.282.

14. Treu, Siegfried — *Testing and Evaluation - Literature Review*, pp.71-88 of "Electronic Handling of Information: Testing and Evaluation". Academic Press, London, 1967.

15. Cleverdon, C.W. — *Identification of Criteria for Evaluation of Operational Information Retrieval Systems.* Cranfield College of Aeronautics, England, November 1964.

16. Buchanan, J.R.; Hutton, F.C. — *Analysis and Automated Handling of Technical Information at the Nuclear Safety Information Center.* American Documentation, Vol.18, No.4, Oak Ridge National Laboratory and Union Carbide Corporation, October 1967.

17. Dugger, E.; Klinger, R.F. — *User Evaluation of Information Services, Information Retrieval: The User's Viewpoint, An Aid to Design.* International Information, Inc., Philadelphia, Pa., 1967.

18. Buchanan, J.R.; Kidd, E.M. — *Development of a Computer System with Console Capability for the Nuclear Safety Information Center.* Proceedings of the ASIS Annual Meeting, Vol.6, Oak Ridge National Laboratory and Union Carbide Corporation, Oct. 1969.

TABLE I

NSIC Budget for FY-69

Operation	*Cost*	*Percent of Budget*
Abstracting and Indexing (8753 abstracts)		
Technical	$ 52,000	13.7
Editing	15,000	3.9
Clerical	30,000	7.9
SUB TOTAL	97,000	25.5
SDI		
Technical	16,000	4.2
Clerical	2,000	0.5
SUB TOTAL	18,000	4.7
Consultation & Retrospective Searches		
Technical	37,000	9.7
Clerical	9,000	2.4
SUB TOTAL	46,000	12.1
Reports and Review Articles	63,000	16.6
Computer Activities	136,000	35.8
Administrative and Other	20,000	5.3
TOTAL	$380,000	100.0

TABLE II

NSIC Costs Using IBM-7090

Operation	*Average Cost*
Input	$ 1.40/document
Indexed Bibliography	$50.00/issue
SDI	$ 0.07/abstract/user
Retrospective Search	$40.00/single search ($25/query in batches)

TA3LE III

Aerospace Research Applications Center Monthly Cost Performance Report

Job	*Section*	*Man No.*	*Hours*	*Base Cost*	*Operations Overhead*	*Total Cost*	*Percent of Job*
RS3342	Engineer	103	8.00	$40.00	$40.00	$80.00	92.8
	Clerical	204	.50	.85	.85	1.70	5.1
	Service	315	.75	2.59	2.59	5.18	2.1
	JOB TOTALS		9.25	$43.44	$43.44	$86.88	100.0

TABLE IV

Aerospace Research Applications Center Cost Distribution of Average Searches

Cost Item	*Retrospective*	*SDI*
Direct Labor	$33.64	$3.22
Phone	4.05	-
Supplies and Document Reproduction	1.83	0.24
Computer Time	27.11	1.38
Operations Overhead	19.43	2.48
TOTAL	$86.06	$7.32

777-01 8 10-09-67

LISTED BELOW IS A SUMMARY OF THE DOCUMENTS SELECTED FOR YOU BY NSIC.
PLEASE CIRCLE ONE OF THE INTEREST CODES FOLLOWING EACH CARD NUMBER.
INTEREST CODE DEFINITIONS ARE -
M - MUCH INTEREST, S - SOME INTEREST, N - NO INTEREST.

CARD NUMBER					CARD NUMBER					CARD NUMBER				
15533	02	M	S	N	16404	03	M	S	N	16430	04	M	S	N
16587	05	M	S	N	17077	[illegible]	M	S	N					

Place Stamp Here

Requires First Class Postage

INSTRUCTIONS

1. The information furnished herewith is believed to be pertinent to your present interests. However, in order to serve you efficiently, we need a clear indication of whether or not this is true. Please acknowledge by checking the appropriate information on the attached "User Feedback Card." Return the pre-addressed card to us promptly to ensure that you receive additional abstracts on similar subjects.
2. When returning the card enter any comments you may wish, such as (1) changes in fields of interest (2) authors of interest (3) change of address, etc.
3. All documents abstracted on these cards may be examined at ORNL, however, the Center is unable to furnish copies of any of the documents with the exception of those with ORNL-NSIC numbers. Available information on document sources is given on the line following the abstract. The addresses of these source organizations are available from NSIC on request.

OAK RIDGE NATIONAL LABORATORY · U.S. ATOMIC ENERGY COMMISSION ·

NSIC

Nuclear Safety Information Center
Oak Ridge National Laboratory
P. O. Box Y
Oak Ridge, Tennessee 37830

Fig. 1 SDI abstract card

NUCLEAR SAFETY INFORMATION CENTER
SDI EVALUATION CHECK LIST

Please complete and return this form to NSIC.

1. I review the SDI abstracts

   ☑ Usually ☐ Occasionally ☐ Rarely

2. I find the SDI service to be

   ☑ Very useful ☐ Fairly useful ☐ Of little use

3. The abstracts are in general

   ☐ Too detailed ☐ Insufficiently detailed ☑ About right

   ☐ Too long ☐ Too short ☑ About right

   ☐ Too old ☐ Too tentative ☑ About right

4. The abstract cards are

   ☐ Retained ☐ Discarded ☑ Circulated to Nuclear Engineers in Company & Filed

   ☐ Filed ☐ Circulated to friends

5. Compared to NSIC's quarterly Indexed Bibliography, I find that

   ☑ I can rely on the SDI service alone

   ☐ The SDI provides too few references

6. In general I prefer

   ☑ SDI ☐ Bibliographies

7. The frequency of SDI mailings is

   ☐ Too often ☐ Too seldom ☑ About right

8. I suggest the following changes in the SDI program:

   None.

Fig. 2 Questionnaire

NUCLEAR SAFETY INFCRMATION CENTE — DEPARTMENT — 10160 — 04 60 70 01 1 — ACTIVITY — DIV 16 — RPT.3057 — 4435-0131 — COST ACCOUNT

FY 69

OPERATING COST REPORT
H.M.BECKLER

		APR		MAY		JUN		FY TO DATE		BUDGET	
		$	MM	$	MM	$	MM	$	MY	$	%
ADM., CLERICAL & OTHER	65L	2221	4.09	2417	4.70			24554	4.10		
FRINGE BENEFIT	69L	437	.89	435	.74			5617	.88		
SCIENTIFIC & TECHNICAL	66L	5350	4.69	5854	4.85			54067	3.78		
SCIEN. & TECH. FRINGE BENEFIT	69L	1104	.82	575	.45			6761	.43		
TOTAL DIVISION LABOR		9112	10.49	9281	10.74			90999	9.19		
DIV. ADM. ON DIV. LABOR		2893		2756				26713			
G & A ON DIVISION LABOR		5667		5796				56249			
LOANEE-IN LABOR	70L	872	.53	1994	1.47			10925	.75		
G & A ON LOANEE-IN LABOR		523		1196				6554			
SUPPORT: P & E SHOPS & MAINT.	50L										
G & A & BURDEN											
I & C ENGR. & SHOPS	50L	18	.01	70	.05			3642	.28		
G & A & BURDEN		25		97				5073			
INSPECTION ENGR.	50L										
G & A & BURDEN											
GE & C	50L										
G & A & BURDEN											
LABOR FROM K-25	50L							321	.06		
G & A & BURDEN								195			
LABOR FROM Y-12	50L							12			
G & A & BURDEN								16			
TOTAL SUPPORT LABOR		18	.01	70	.05			3975	.34		
G & A & BURDEN		25		97				5284			
SHOPS & MAINTENANCE	50M							7			
SUBCONTRACT	61M	24		56				753			
TRAVEL	62M			298				1455			
MAJOR PROCUREMENT	64M							177			
MISC. MAT'L. CHGS. & CREDITS	65M	4514-		1012				17474			
EDP (CDP & ORNL)	73M	11065		12165				125353			
WORKED MAT'L. & OTHER CHARGES											
TOTAL MATERIAL		6575		13531				145219			
TECH. SERVICES: UTILITIES											
ANAL. CHEMISTRY											
HEALTH PHYSICS											
ORR											
BULK SHIELDING											
HOT CELLS											
TANK FARMS											
M & C LABS											
TECH. INFORMATION DIV.											
TOTAL GROSS COST		25685	11.03	34721	12.26			345918	10.28		
DISTRIBUTION OUT											
INVENTORY CHANGE											
NET OPERATING COST		25685		34721				345918			
MEMO: OT FOR MONTHLY PERS.			.12		.13				.07		

Fig.3 Operating cost report

SUMMING-UP OF THE SPECIALIST MEETING

Melvin S.Day

Acting Assistant Administrator for Technology Utilization
National Aeronautics and Space Administration
Washington, USA

I believe that all of us have been impressed by the dominant theme of this meeting: practicality, combined with common sense, in meeting the problems of mechanizing an information center. If it has been true in the past that, as Mr Isotta aptly indicated, often times "a program of mechanization was initiated and pursued simply because it was the thing to do", this is no longer so. Today we fully realize that automation is essential and must be accelerated, but we are also conscious of the associated problems. Many of these problems have been discussed at this meeting. Among them are the continuous advances in computer capabilities, sources of input to the automated systems, availability of full text of the original documents, selling top management on the need for a computerized system, difficulties of training users to benefit from the new system, and system management.

Yesterday morning Mr Jones opened our meeting with a down-to-earth hard hitting approach to the critical problem of obtaining management approval of the proposal to mechanize. His experience confirms that there is no substitute for painstaking preparatory work and a well-presented briefing. He also stressed that the terms used in the presentation must be the same terms that are used by management in its evaluation of other branches of the organization. I feel that this is a particularly important point with respect to computer applications. Management has certainly not become disenchanted with computer operations - far from it - but management has become more concerned with the proper place of the computer in the organization. It has become familiar with systems analysis, and is conscious of the precautions that must be taken to assure operational success. Therefore, Mr Jones' advice not to oversell, and to be ready for a supplemental study to cover other points that management wants, is particularly appropriate.

Mr McCarn reinforced this need for careful planning in readying an organization for automation by telling a horror story of a system failure and the fortunately happier story of the successful development of a MEDLARS system. We should all be conscious of the conflicts that will develop in implementing a new system and attempt to resolve as many as possible in the planning stage. As Mr McCarn points out, there may be realistic objections to the proposed new system. Probably all of us are familiar with unjustified enthusiasms for automated systems that prove to be impractical because of neglected aspects of design and planning. So here again, emphasis should be placed on design and planning.

Specific problems in transfer to computer operations were represented by Mr Amey's paper on finding the optimum mix of staff and consultants and the effect of computer hardware on implementation and performance. Mr Amey noted that some systems characteristically provide services that are computer-oriented rather than directed toward the needs of the ultimate users. I am sure that we all know of black box systems with input and output not adequately fitted to human users and difficult to expand or transfer to a different make of computer. Involving your internal staff heavily in the design of the system is difficult but desirable, as Mr Amey implies. Unless your staff is forceful and has adequate expertise in computer capabilities, you will be amazed at the power of the computer to force your planning into its limitations.

Mr Schuler's paper gives many practical aspects of a real life information system including procurement of computer equipment, and recruitment and training of staff. In his paper he also illustrated path chart approach.

Mr Dammers forcefully brought one of our difficulties to the front: computer development appears at first to outrun our capability to apply it to our best advantage. Continuous rapid advances in computer speed, memory size, and peripheral equipment, combined with variables of delivery date, cost performance, and compatibility with existing equipment, make hardware selection a compromise. Nevertheless, in planning ahead, many information systems may expect to parallel the developments planned by Mr Dammers for his own system. In particular, on-line mass storage, real-time rapid response query systems, and visual display extended to many access points, appear at this time to be on the path we shall almost inevitably follow. Direct access to a large computer, or use of a satellite computer as described by Mr Dammers, or remote terminal connection to a commercial information service are possible system alternatives.

The need for operators of small information centers to understand systems design, even if they are dependent on larger centers, was presented by Mr Bree. In particular, he stressed the need for a controlled vocabulary; that is, a thesaurus. This too seems to be a firm trend. With apologies to Mr Dammers, I say that I feel the path to the use of a completely free vocabulary does not appear likely to be well traveled. So we shall have to develop ways of permitting the use of an information system to use a thesaurus conveniently. Mr Bree has

described the graphical form in which the EURATOM Thesaurus is presented. Here again direct access, remote-terminal computer systems could permit thesaurus terms to be directly displayed on a cathode ray screen. The user can then browse through the thesaurus while setting up his search query. He is also informed of the number of documents that would result from any stage of his search before a display of the citations is requested by the browser.

The user's ability to formulate his own search may change the present pattern in which a documentalist acts as intermediary between the scientist and the system. Mr Vernimb discussed the problem of the scientist who attempts to use an information system by a single-term approach and described a successful attempt to get more specific inquiries for a "batch" type processing system. Here again, the remote terminal computer system could provide the "instant feedback" of which Mr Vernimb speaks.

Besides the technical developments in information science and in computer technology which we must consider, costs are becoming everyone's concern. As we are all aware, today's research and development programs and the information services to support them exist in an atmosphere of stringent economy. All budgets are undergoing close scrutiny. Because of the expense of computer operations, management will insist on meaningful cost and effectiveness data. In his paper on evaluation and cost analysis of automated information analysis center operations, Mr Buchanan has shown how these data can be collected. He has stressed that the essential ingredients are to plan early for data collection, and to design a system that allows for easy interpretation of results. To turn from the computer aspect of automation, an essential ingredient of an information system is the provision of the full text of desired documents. If these have been located by fast computerized search, the user will also expect to receive the document quickly, as Mr Amey emphasized. This presents acquisition problems to the information center. Should complete sets of documents be purchased, or should only those in certain categories be ordered? If the latter, individual documents not in these categories will have to be ordered at increased unit cost and some delay.

Furthermore, a significant trend in the US affecting costs of information is the termination by several large US Government agencies of free distribution of documents to their own contractors, universities, and other agencies. Last year, the Defense Documentation Center began to charge its users for its unclassified documents, and this year the Atomic Energy Commission and NASA have followed suit in introducing user charges for many of their information products. Information centers will therefore have to consider the effects of these charges and provide for them in their document acquisition budgets. This will probably lead to a greater emphasis on the use of the less expensive microfiche. While microfiche are inexpensive in themselves, the information center will have to consider the costs of readers and printers, and possibly microfiche reproducers. An optimum distribution of costs in acquiring the various equipment should be determined on the basis of real user needs. Before closing, and with your permission I would like to draw upon my own NASA experience with our aerospace information for some final comments. Although the NASA scientific and technical information system must be considered large by any standard, the problems encountered in automation and the trends that have developed are applicable, in my opinion, to any sized information operation. Mr Isotta described in an excellent paper the adaptation of this system to a European environment.

In many cases, our experiences have exactly paralleled those described at this meeting. In computer operations, we began with an IBM 1401 in 1962, and replaced it with a 1410 less than two years later. In 1966, an IBM 360 Model 40 was installed, last year an IBM 36050 was added, and we are just now phasing out Autocoder programs that have been run in emulator mode. This is a good example of the impact of a particular piece of hardware configuration that Mr Dammers mentioned this morning.

We very early made plans to avoid being limited by software; that is, to get away from the problem of re-writing every complicated computer program whenever new equipment becomes available. A high-level programming language, PL-1, was selected, and almost all our programs are written in this. Our publication system is table driven, so that reformatting a print program, for example, is greatly facilitated.

For our indexing and search programs, we took the path of a controlled vocabulary and prepared a NASA Thesaurus. This is compatible with the LEX Thesaurus mentioned by Mr Schuler and is consistent with Mr Bree's conviction for the need of a controlled vocabulary.

Computer printing has been avoided in favor of photocomposition wherever possible. Our abstract journal *STAR (Scientific and Technical Aerospace Reports)* is prepared by a combination of photocomposition methods. The abstract section is composed on computer-tape-driven Photon equipment, while the indexes are set up on the computer-driven Linotron.

Current awareness service is provided to NASA Centers and contractors. Initially, an SDI program using individually tailored interest profiles was developed. We operated this system for 6 years. At one time there were about a thousand individuals participating in our SDI program. In view of the availability of an alternative current awareness system known as SCAN (Selected Current Aerospace Notices), SDI has been terminated. SCAN provides computer-selected notifications under 186 subject topics. Because the single computer-printed listing is reproduced by low cost offset printing, SCAN cost per user goes down with increasing numbers of participants, typically being one-fifth to one-tenth the cost per user of SDI. We now have almost 9000 individuals receiving one or more SCAN notification each month. You may recall that the success of this system in the US has been duplicated by ESRO's success in Europe. Mr Isotta mentioned ESRO's corresponding successful Standard Profile Service.

NASA's RECON, short for REmote CONsole, is the program which above all, we believe represents the path of the future for information centers and users. Although it is still in the early stages of installation, we shall have CRT display terminals in all our major NASA research centers by November. The RECON user will first see appropriate sections of the NASA Thesaurus displayed. He can then formulate his search by inserting Boolean relations between Thesaurus terms, and be informed at each step of the number of documents citations that the search would reveal. Bibliographic citations of the selected documents are then displayed or printed out as the user wishes.

Besides RECON use in retrospective searches, we plan to have a current awareness capability operational next year. To see citations to documents that have entered the system since the user's last contact with RECON, he will need only identify himself, and the citations selected by his personal SDI interest profile or by desired SCAN subject topics will be displayed. RECON is representative of the many direct access information systems that will be available to numerous information centers in the future, either through in-house operation or through terminals connected to commercial systems.

Mr Chairman, I am delighted to have been able to attend these informative sessions. By sharing our experiences as we have done here, I believe that we are helping to chart out the course that developments in information science and technology will take, over the coming years. The papers presented at this meeting are first class. In themselves they are a meaningful contribution to the state-of-the-art literature. I salute each of the speakers and to each of them I express both my admiration and my thanks.

ADDRESS OF APPRECIATION

G.H.Tidy

Defence Research Board, Ottawa, Canada
AGARD National Delegate

It gives me great pleasure to express for Canada our appreciation for being given the privilege of holding this Specialist Meeting on "The Mechanization of Small Information Centers" here in Ottawa. From the papers I have heard during this meeting, I am encouraged to believe that substantial progress can and will be made toward increased efficiency and sophistication of small information systems at reasonable cost.

I am particularly encouraged by the way that our librarians, accustomed in the past to working in the quiet atmosphere of booklined libraries, have not hesitated to adapt to the computer age, in spite of the possible disruption of the traditional quiet of their sanctum by the noisy clatter of the ancillary equipment! One could, of course, also interpret the union of automation with information handling as being initiated by aggressive computer specialists!

However the union may have taken place, there is no doubt that it can be most fruitful in improving our ability to keep pace with the ever increasing flood of technical information. It is most important for our technological future that the output of scientific and engineering knowledge so necessary for progress be retrievable selectively and quickly.

This conference has indicated that many problems remain to be solved but I am convinced that with the skills and energy being devoted by you, substantial progress will continue. I hope that the conference has assisted in this.

In closing I wish to express my appreciation to the Executive Committee of the Panel for accepting Canada's invitation, and I thank particularly the speakers and the behind-the-scenes workers who have contributed to much to the success of the meeting.

DISCUSSIONS

The discussions which followed the presentation of the papers were kindly compiled from the tape recording by Mr J.J. Irvine from the D ence Research Board, Ottawa.

PAPER No. 1 - Mr A.C. Jones

Mr Buchanan, USA: Why is it so important to get the management to realize the complexity of your information work?

Mr A.C. Jones: There are two reasons. One is that the great advantage of computers is their ability to solve problems that cannot be handled readily in any other way. Therefore, if you cannot demonstrate that your present system is fairly complex, there is not much hope of getting approval to change it.

The main reason is that the complexity of the system determines the quality of the people you are allowed to employ.

Dr Vernimb, Germany: You referred to the need to prepare quantitative data about your present system. Is this not expensive to obtain?

Mr A.C. Jones: Yes indeed. In a manual system they are expensive. If people are gathering statistics, they cannot be doing their usual jobs. But, if you don't get the statistics, there isn't much hope of demonstrating the need for the computer.

Mr McCarn, USA: It seemed to me from your description that most of the motivations you were appealing to in management were concerned with improvements in your organization and the service it could provide. Do you not also appeal to other motives such as their own esteem as managers of the system or to nationalistic aspirations? Could you comment on other motives?

Mr A.C. Jones: Yes, I think both of the ones you have suggested are real and significant. Management is obviously desirous of seeing every part of its organization improved. It is always nice to feel that you are keeping as up-to-date as possible within your resources. Possibly the greatest of the secondary motives is the hope that the new system will result in a body of satisfied customers.

Mr McBurney, Canada: You will recognize this as a dirty question and I do not really expect an answer. In all levels of the government there is pressure to measure the cost-benefit ratio of services. This is probably more difficult to do in defence work than in the civilian sector, but do you have any magic method of measuring the benefits to your users of your services?

Mr A.C. Jones: I am glad that you said that you didn't expect an answer. We collect letters of appreciation, but have as yet developed no quantitative measurements. Is there anyone in the audience who would care to comment on this problem?

Dr Wooster, USA: The technique of collecting testimonials on what is sometimes called a Gabriel file is quite widespread. I have recently done an evaluation for the US Committee on Scientific and Technical Information of the literature on measuring the effectiveness of information analysis centers. The evaluation will be published in American Documentation in the near future. The conclusion was that there is no satisfactory standard of measurement.

Dr Campbell, Canada: In preforming your study and presentation to management, how much user participation did you require and encourage?

Mr A.C. Jones: If you mean user participation in developing the proposal and presenting it, the answer is very little, possibly because of the lack of users in close proximity. What I personally would have liked to do would have been to encourage direct contact between the users and management. For the future, we have established a steering committee on which the users are heavily represented and from which we hope to get useful guidance.

PAPER No. 2 - Mr S.C. Schuler

Mr Morgan, Canada: Mention was made in a slide of the provision of means of eliminating documents of low scientific value. Could you explain the criteria used in this selection and who does it? Does the user have an input in this? Also, what percentage are discarded?

Mr S.C. Schuler: The user doesn't have much to say about this directly. The selection is made by the subject specialists on the basis of their experience of what is useful. It should be emphasized that some material in our collection does not receive the full treatment. For instance, in our journal we have a section that contains subject listed titles only.

Miss Marriott, Canada: What kind of people do the editing of your tapes? Are they clerical or professional?

Mr S.C. Schuler: The checking of the tape is done by the supervisor of clerical staff, but the technical content is checked by an editor before it gets on tape.

Mr Gillan, Canada: You made mention of dedicated positions on the paper tape. Will this cause problems in the choice of number of tracks in computing equipment later on?

Mr S.C. Schuler: We are using seven track tape and the size of the file we will have built up by the date of installation of an in-house computer will make it difficult to change to another standard.

Mr Atkins, Canada: You mentioned problems of legibility of computer output. Have you made arrangements to check this quality by your organization prior to press run?

Mr S.C. Schuler: By liaison between our organization, the computer operators, and the printer, most of these problems have been eliminated and we are now satisfied with the output.

Mr Jones, Canada: Could you explain the reasons for your plans to move from a service bureau computer to an in-house computer?

Mr S.C. Schuler: The HMSO computer is presently used by many other organizations and it will not have capacity for other operations which we will wish to add in the future.

Mr Day, USA: Could you tell us more about the Tech Link Service? How many users do you have and what use do they make of the Tech Links?

Mr S.C. Schuler: We currently have about 4700 users in industry. This service is not advertised, but rather is distributed through nine regional offices. Each regional office is managed by a research scientist and is responsible for collecting the area's requirement for Tech Links. The user population is growing at about sixty per month and about five new Tech Link items are prepared each week.

We have no way as yet of quantifying the utilization of the Tech Links. One notable success was signalled by the request for 1500 copies of an item on optimum routing of delivery vans.

Dr Wooster, USA: Have you thought of changing from a loan service to a system of giving away or selling reports?

Mr S.C. Schuler: We do sell reports listed in Tech Links. The loan service of other reports is presently included in the subscription price of the announcement journal. Plans are being made to charge directly for the loan service of paper documents in 1970. We are actively encouraging the use of microfiche, but their widespread adoption is dependent in the availability of better and cheaper readers.

PAPER No. 3 - Mr N.E.C. Isotta

Mr Atkins, Canada: In your preparation for information service in the field of electronic components, did you distinguish between state-of-the-art and state-of-the-market?

Mr N.E.C. Isotta: Yes. There is greater emphasis on state-of-the-market because even the designers of future spacecraft are reluctant to specify components that are not presently available.

Dr Brown, Canada: I presume your services are limited to ESRO member states. I am also curious about the division of use of your services between ESRO establishments and outside organizations.

Mr N.E.C. Isotta: The service is limited to ESRO countries. It was originally set up for ESRO staff, currently comprising about 1200 people scattered throughout the member countries. At the moment about thirty percent of reutilization of the service is by staff members. However, many of the outside users are in institutions who are developing experiments for the space craft. In this sense, these institutions are virtually considered as ESRO establishments and get the same information service.

Mr McGee, Canada: I am intrigued by your customer liaison group. Can you say something about their opinions of the descriptor problem of picking the right word to retrieve the right document?

Mr N.E.C.Isotta: Of course there is a big problem and it is probably worse in Europe because of language difficulties. Retrieval calls for considerable familiarity with the thesaurus and I tend to think that the people who are retrieving should also have experience in indexing the documents. Secondly, you cannot expect these people to have expertise in a wide range of subjects. They have to make the best selection of terms they can, from a fairly narrow speciality base by comparison with the coverage of the NASA system.

Another difficulty is inherent in the indexing philosophy on which the thesaurus is based. We index to the most specific term and this entails a careful search of related thesaurus terms when a broad-based search is undertaken.

A point that should be made is that the situation is vastly changed with the introduction of NASA's RECON. Here, the interaction between user and computer greatly facilitates the formulation of the search.

Dr Cherrington, Canada: How much noise will the user tolerate in the information system?

Mr N.E.C.Isotta: In the case of SDI, I would say up to one hundred percent which may be a controversial statement. In a data base of one thousand, the user might well be willing to pay for twenty announcements per month when only ten are relevant. In a larger data base this percentage would not of course be tolerable.

Mr Dammers, UK: Could you give us data on the cost of your information service in ESRO?

Mr N.E.C.Isotta: It is about one and one-half percent of the ESRO budget which is in line with accepted standards.

Mr Willis, USA: Would you give us some idea of the number of your staff in the information center in Paris and in the computer center in Germany?

Mr N.E.C.Isotta: This is rather difficult because our work is only part of the load on the computer. We have about one and one-half system engineers and a documentalist in Germany together with one system engineer in Paris. In addition there are a further twenty-two staff based in Paris.

PAPER No.4 – Mr D.B.McCarn

Mr Hinckley, UK: Do you have any other concrete suggestions to system developers to obviate this hostility from staff and management?

Mr D.B.McCarn: Part of the answer is in the attitude of the computer people. If they think of themselves in the Messianic role of solving other people's problems by a solution imposed from on top, then almost certainly there is going to be trouble. If they see their roles as co-operating in the development of a system, they are much more likely to have a successful development.

Dr Christie, Canada: Can you pinpoint any reason for the lack of information on the problems of selling an automated system to management?

Mr D.B.McCarn: The automation people are more interested in the technical problems of development than in human problems. They tend to be concrete-minded rather than being interested in people.

Mr Dammers, UK: You said that a stable organization opposes innovation. How about a system that is growing?

Mr D.B.McCarn: If a growing system can see a possibility of handling the increasing load, then you will have difficulty in automating the system. If the organization cannot handle a growing work load, it is unstable.

PAPER No.5 – Dr G.X.Amey

Dr Wooster, USA: When you mention consultants in your paper, are you referring to your contractors?

Dr G.X.Amey: Yes.

Dr Wooster, USA: What is your opinion of the COSATI subject headings for classification of reports?

Dr G.X.Amey: There are difficulties in its use but it is probably as good as any other classification. Most classification systems are essentially symbolic so if one is consistent in its use, the choice of system doesn't make too much difference. We also use the TEST Thesaurus, but our impression is that it is gaining wide acceptance except in the US DoD which was the developer.

Dr Vernimb, Germany: Is it really necessary to develop your own system? Are there not systems available which could be used?

Dr G.X.Amey: Perhaps this will be possible in the future, but at present there is a lack of standardization in computers and software. There is a concept of software components which might be available in the future, but if you want a system now, it is necessary to build your own.

PAPER No.6 - Mr H.F.Dammers

Dr Wooster, USA: Do you plan to install graphic terminals in your system as well as the present teletype terminals?

Mr H.F.Dammers: We are using CRT terminals now mainly as a means of dispensing with multiple remote card indexes. As for light pens and reactive graphic terminals, the problem is one of cost and we have no immediate plans to install any.

Mr Booth, Canada: I am interested in your card system. It seems to be somewhat like the Termatrex systems which are familiar on this continent. These are often abandoned in favour of computer systems when financially feasible, but you run both systems in parallel. Is this to provide access when remote terminals are not available?

Mr H.F.Dammers: Yes, exactly. The card systems are an extremely convenient way of providing access to the storage when remote terminals cannot be provided for economic or hardware reasons.

Mr Bree, Germany: There is clearly a trend towards large input processing centers which would provide a service to smaller information centers. There are two possible means of co-operation between the two types of centers, direct on-line connection or independent operation of the smaller centers using magnetic tape material supplied by the input centers. Could you comment on the choice between these modes of operation?

Mr H.F.Dammers: My philosophy is to avoid doing anything that is available elsewhere in a cheaper or better form. Our center has an obligation to process data produced by our company, but we see no reason to handle published literature. We will gladly take whatever is offered by Chemical Abstracts and other similar services. This holds true for companies and I think will be also true for National organizations.

Miss Evans, Canada: Can you give us figures on the proportion of equipment cost to system cost?

Mr H.F.Dammers: In our case, it is about thirty percent which is higher than usual. I think a more common figure would be twenty percent.

Dr Amey, Canada: Could you say something about budgeting and staffing problems in your evolutionary system? Did you hire supplementary staff in the conversion period or get additional funds?

Mr H.F.Dammers: Any increases were as a result of the expanding flow of information and not specifically for mechanization. We have a continuing policy of allotting twenty-five percent of available effort to development. You might ask how this can be done in a busy library where there is no money or effort available for mechanization. Well, information is a flexible business and usually money can be found by slight reductions in purchases or services.

Our fraction of the research budget has declined from three percent in 1962 to two percent in 1969.

Mr Willis, USA: How do you estimate the useful life of your equipment?

Mr H.F.Dammers: This is difficult because of the rapid developments in the field and the long lead time in procurements. The problem is to decide whether to buy or rent equipment. If one keeps a computer for five years it is cheaper to buy, but perhaps the computer will no longer be suitable at the time.

Perhaps the best plan is to buy equipment that can be integrated with future purchases. For instance, our present computer will become a slave unit for the next model purchased.

Mr Morgan, Canada: Does the information service's two percent share of the research budget include computer and programming costs?

Mr H.F.Dammers: The budget figure includes computer time and the development of the programs.

PAPER No.7 - Mr R.Bree

Mr McGee, Canada: I run a small information system and most of my decisions on policy have been the opposite of yours. Do you think that larger centers are mission-oriented or function-based while smaller centers are subject-based?

Mr R.Bree: I presume that your reference to a small center is to one that has a limited subject field. In that case the subject cataloguing can be quite detailed while in a larger mission-oriented system the subject coverage is necessarily more superficial. The decisions on policy are somewhat different for the two cases.

Dr Wooster, USA: A comment first. There are two manufacturers providing equipment to store and retrieve micro-fiche under computer control.

My question has to do with the meaning of weighting as used in your paper when referring to search techniques. What do you mean by the term "weighting"?

Mr R. Bree: It means printing of the citations of most likely interest at the beginning of the retrieval list.

PAPER No. 8 – Dr J.E. Brown

Mr Laval DuBeuil, Canada: Do you intend to run retrospective searches and if so when?

Dr J.E. Brown: Yes and no. The tape services haven't been going long enough to provide experience on retro-spective searching but we are prepared to extend our service from current awareness to searching when the demand is sufficient. One important point in current awareness service that we recognized was the necessity of ensuring that we held all journals that the tape services referenced. We can provide photocopies of any articles to which the customer may not otherwise have access. A charge is made for photocopies, but the journal may be borrowed on inter-library loan.

Mr Dale, UK: Does your flat charge of one hundred dollars apply to all your customers and what do they get for it?

Dr J.E. Brown: The one hundred dollars is a flat rate which theoretically enables the customer to use all three tape services to which we subscribe. In practical experience, most customers find that just one of the tape services provides an adequate coverage in his field of interest. The charge covers the development of a profile of up to sixty terms. There is an increased charge for more than sixty terms.

One other point about the tape services is that they are prepared from advanced copies of the journal articles so we sometimes have the embarrassment of announcing the article in current awareness before the journal is available.

Mr Dagger, Canada: Do you find that your inter-library loan service is impeded because of incomplete information in the requests?

Dr J.E. Brown: Yes. We have done a study that showed that sixty percent of the time spent on this service was due to inadequate or ambiguous bibliographic details on the requests.

Mr Dagger, Canada: What percentage of customers are satisfied with the relevance of the current awareness service?

Dr J.E. Brown: There is a tremendous variation in opinions about importance of relevance. Some scientists are content with twenty percent and others expect and profess eighty percent. There is difficulty in getting the customers to keep their profiles up to date.

PAPER No. 9 – Mr J.H. Klopp

(In the absence of Mr Klopp, his paper was read by Mr Stolk. He suggested that questions should be written out and mailed to the author.)

Mr Dammers, UK: Unless Mr Klopp had an enormous system to develop and set up, which wasn't quite apparent from the task which he outlined, I wonder how he could afford so much pre-training? I am not quite clear whether he is reporting or proposing.

Mr Stolk (Partly inaudible): From what I understand, only part of this was actually done so I think that the paper was a picture of how it should be done according to him.

PAPER No. 10 – Dr Vernimb

Mr Isotta, ESRO: This is not so much a question as a comment. I would like to say how much I agree with the speaker's evaluation as a most desirable one for an interactive operation. I think it is worth mentioning that the NASA RECON does almost exactly what he has described.

Dr Vernimb: I agree.

Mr Dammers, UK: The difficulties which Dr Vernimb indicated, when a user tries to query a system using a specified indexing language, are to be expected. If one translates the text via a special indexing language, the user gets into difficulties and another complicated system must be added. Would it not have been easier to forget the indexing and try to live with free language?

Dr Vernimb: This is certainly a possibility and there are systems which compare the text of the queries with the text stored in the system. But in my opinion we have to use a structured thesaurus or vocabulary in order to get an acceptable recall ratio and because it is a cheaper system. It is partly a question of economics.

Mr Dammers, UK: I do not think that there is any proof that a system based on indexing achieves a better recall ratio than one using free language. On the contrary, whatever trials have been done have shown only marginal differences.

With regard to costs, I would like to know if you have specific information on which to choose one system over the other.

Dr Vernimb: In a free language system, much computer time is needed, whereas in a vocabulary system, comparatively little computer time is required. This is the reason we have not shifted over.

Dr McIvor, Canada: In our experience with chemical abstracts where we are using free text, I find that it is not unreasonable to expect the users to prepare their own queries with some encouragement. I agree that it does take considerably more computer time to search.

Are these terms that are used for the search prepared by the same documentalists that are preparing the profiles? Have you considered the time used by the documentalists as well as the computer time in the total system cost?

Dr Vernimb: Our relatively small team of specialists started with indexing but it was found necessary to contract out this work. The in-house team checks the outside indexing and also prepares the profiles and queries.

Mme Roeper, France: Ce n'est pas exactement une question mais j'aimerais revenir sur ce qu'a dit Dr Vernimb. A mon sens on passe le temps soit à l'entrée en ordinateur en analysant soigneusement les documents et on récupère ce temps à la sortie ou alors on choisit la méthode inverse, et à mon sens les deux se rejoignent à moins qu'on entre simplement des documents et qu'on veuille jamais répondre aux questions.

Dr Vernimb: The economy of this procedure is a function of the number of documents entered and the number of queries asked. In our system with a yearly input of 100,000-120,000 documents, we use about eight to ten minutes for indexing and about half an hour for query formulation. We handle about one thousand queries per year with our small staff and we think this is not too bad from an economic point of view. With the relatively small indexing time, we think that it is an economic system.

Mr Dammers, UK: As the last question was relevant to my question, I would like to pursue it further. I think it is wrong to evaluate the system simply on the basis of indexing time as this is usually underestimated. I find that the apparent time of indexing and abstracting does not account for the full time of the staff. Apparently more time per document should be allocated for a realistic accounting. But more importantly, one should look at the costs of the total system and this should include the user as well. This is very difficult to quantify but one should consider the delays suffered by the user and how easily the user interacts with the system.

With regard to free text, very few of us would say that one should put in full text that which is still impractical. Rather, words as they occur in the text should be freely chosen and the point is that they are the natural text words and are not translated into words which have fixed and determined meanings. Vocabulary systems work all right as long as the indexers also formulate the questions. This means that the translation and deformation occurring in input is allowed for at the output but is very difficult to guarantee in a large organization.

Dr Vernimb: As for your first comment, certainly indexing cannot be done full time. Our contract indexers do this work in their evenings and weekends as an addition to their regular work in the front line of research in their speciality. So they have the best knowledge in the field and are well equipped for indexing after some training in documentation.

PAPER No. 11 - Mr J.R. Buchanan

Miss Ironside, Canada: You stressed the importance of using technical specialists to do the subject cataloguing of your material. Would you enlarge on this point, are they employed full time on this and how do you meet the degree of specialization required?

Mr J.R. Buchanan: They are bench scientists and engineers who are active in their field and for this reason we are content to put up with some of the problems inherent in dealing with part time staff. The clerical staff is, of course, full time.

Dr McIvor, Canada: Could you tell us if your system is based on individual profiles and, if so, are they prepared by the user or by specialists? Is it based on a thesaurus or on free text searching of the abstract?

Mr J.R.Buchanan: We have a thesaurus of seventeen hundred terms. About one third of the profiles are based on thesaurus terms. The query comes in prose form and a profile is developed from thesaurus terms by our specialists. The remainder of the profiles are based on standard categories. Our subject field is divided into twenty categories and many users find that their interests fit into these categories. There is provision for feedback and consequent negating of some terms in these categories.

Also, we have developed specialized profiles that cater to large groups of users such as the instrumentation staffs of reactor power plants.

Mr Irvine, Canada: What elements of information are stored on your computer and how is it inputted?

Mr J.R.Buchanan: We store the usual bibliographic information – the author, corporate author, miscellaneous information on number of pages and figures. We also store the key words or terms from the thesaurus and a one-hundred word abstract.

Until about three months ago this information was inputted to the computer by first being typed on a conventional typewriter at our own facility. This was then sent to the computer facility and went through the key punching and then was processed into our master tape which now contains about 35,000 items. So we are essentially a small mechanized information system.

But starting in June we converted to the direct input system and we are now inputting via a CRT device so that all of our references are now being entered, proofed, and corrected on the CRT. They are then dumped into the main file and are immediately retrievable by remote terminals and by any retrospective searches that may be run at the computer facility.